DOCUMENT RESUME

ED 214 957 TM 820 108

AUTHOR Quellmalz, Edys; Snidman, Nancy
TITLE Evaluation Workshop III: Program and Evaluation Planning. Participant Notebook. New Edition.
INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.
PUB DATE [82]
NOTE 193p.; For related documents, see TM 820 106-111.

EDRS PRICE MF01/PC08 Plus Postage.
DESCRIPTORS Data Collection; *Decision Making; Elementary Secondary Education; *Measurement Techniques; *Program Development; *Program Evaluation; *Training Methods; *Workshops

ABSTRACT

A Center for the Study of Evaluation (CSE) evaluation framework describes the decision to be made, appropriate evaluation activity and activity description. The topic of this third CSE Evaluation Workshop is program planning, the second stage of the four stage evaluation framework. Evaluation activities should follow an orderly and logical sequence beginning with Needs Assessment, through Program Planning, Formative Evaluation and concluding with Summative Evaluation. In practice, an evaluator rarely participates in all four stages of evaluation for a single program. Practice exercises are provided for workshop participants. Evaluation questions must be identified and them means by which they will be answered must be selected. Data collection instruments should be carefully selected. Measurement instruments are chosen in accordance with a program's objectives. These may be cognitive and psychomotor or affective. Measurement devices may be tests, reports, observations, or records. Successful completion of this workshop entitles participants to conduct workshops in evaluation for other groups. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214957

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
[X] This document has been reproduced as received from the person or organization originating it.
[ ] Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

EVALUATION WORKSHOP III
PROGRAM AND EVALUATION PLANNING*

Project Director: Edys Quellmalz

Project Coordinator: Nancy Snidman

Developers: Marcella Pitts
Kathi Yohn

Writers: Darrell Couturier
Rachel Grala
Russell Hunter

Editorial Assistant: Shelby Popham

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

G. Gray

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

*A new edition of the third of the CSE Evaluation Workshop series. Most material is orignal; some sections are adapted from the previous version.

EVALUATION WORKSHOP III:

PROGRAM AND EVALUATION PLANNING

PARTICIPANT NOTEBOOK

CONTENTS

Section		
A.	Introduction to the Workshop	
	Goals and Objectives	A - 1
	CSE Evaluation Framework	A - 2-5
	Exercise # 1: Categorizing Evaluation Activities	A - 6
	Feedback	L - 1
B.	Introduction to the Simulation	B - 1-5
C.	Overview of Program Planning	C - 1-4
	Exercise # 2: Task and Responsibilities in Program Planning	C - 5
	Feedback	L - 2
D.	The Program Planning and Evaluation Context	D - 1-5
	Exercise #3: Program Planning and Evaluation Context	D - 11-25
	Documents File	D - 6-10
	Feedback	L - 3-6
E.	Program Planning	
	Introduction to Program Planning	E - 1
	Clarifying Objectives	E - 1-5
	Developing the Program	E - 6-7
	Selecting a Program	E - 7-8
	Program Selecting Techniques	E - 8-24
	Exercise # 4: Choosing a Program	E - 25-33
	Feedback	L - 7
F.	Building in the Information Plan	
	Evaluation Questions	F - 1-7
	Formulating Evaluation Questions	F - 7-9
	Exercise # 5: Formulating Evaluation Questions	F - 10-11
	Feedback	L - 8-9

G.	Measurement Considerations	G - 1-7
H.	Choosing Tests to Measure Achievement Outcomes	H - 1-9
I.	Measurement Devices	I - 1-7
	Measurement Concerns for Evaluation Questions	I - 8-14
	Planning an Efficient Measurement Strategy	I - 15
	Exercise # 6: Developing Measurement Devices	I - 16-19
	Feedback	L - 10-12
J.	Evaluation Design and Analysis	
	Evaluation Design	J - 1-8
	Analysis	J - 8-17
	Exercise # 7: Determining Appropriate Design and Analysis	J - 18-23
	Feedback	L - 13-17
K.	Written Program Plan	K -1-2
	Exercise # 8: Summary of the Model Program Plan	K - 3-10
	Feedback	L - 18
L.	Feedback to all Exercises	
	Exercise # 1: Categorizing Evaluation Activities	L - 1
	Exercise # 2: Task and Responsibilities for Program Planning	L - 2
	Exercise # 3: Program Planning and Evaluation Context	L - 3-6
	Exercise # 4: Choosing a Program	L - 7
	Exercise # 5: Formulating Evaluation Questions	L - 8-9
	Exercise # 6: Developing Measurement Devices	L - 10-12
	Exercise # 7: Determining Appropriate Evaluation Design and Analysis	L - 13-17
	Exercise # 8: Summary of the Model Program Plan	L - 18

SECTION A

INTRODUCTION TO THE WORKSHOP

Section A

GOALS

A. Introduction to the Workshop
   1. Goals
   2. Objectives
   3. CSE Evaluation Framework
   4. Exercise 1

Evaluation Workshop III: Program Planning Evaluation is designed to enable you to acquire evaluation skills that are essential in educational program planning. The evaluator's main responsibility at any stage of an educational evaluation is to provide information to decision makers. In program planning, the evaluator must have the skills to:

1. Help planners develop and implement procedures for identifying alternative programs and for selecting among them.

2. Build the evaluation plan into the program plan.

A participant who completes Evaluation Workshop III: Program Planning Evaluation is eligible to lead the workshop for another group. A leader should, of course, have experience in the areas covered by workshop, and should view it as one means of communicating expertise. If you are interested in conducting future sessions of our workshop, please discuss this matter with your present workshop leader.

OBJECTIVES

Evaluation Workshop III: Program Planning Evaluation is not designed to train people to develop programs, nor to provide all the training necessary to construct the entire evaluation design for any given program. Specifically, the workshop objectives are to train participants to:

1. Distinguish the major characteristics of the CSE framework for evaluation, and be able to apply them in a particular setting.

2. Identify appropriate roles and responsibilities for the evaluator and other people working within the context of program planning.

3. Identify appropriate instruments for helping to select programs to meet particular needs.

4. Select the information system for conducting program evaluations to include:

   Design and sample to be used

   Measurement instruments

   Data collection and analysis methods

   Scheduling of evaluation tasks

   Identify techniques for selecting or developing appropriate evaluation measures.

5. Critique evaluation plans and their appropriateness for a particular context.

CSE EVALUATION FRAMEWORK

Program planning, the topic of this workshop, is the second stage of the four-stage evaluation framework developed by CSE. Supporting the evaluation framework is one key notion: the purpose of evaluation is to aid the process of educational policy and decision making. With this purpose in mind, evaluators have the twin responsibilities of securing trustworthy information about an educational enterprise and conveying the information -- in a timely and useful form -- to people whose actions will be guided by the evaluation effort.

Evaluation, as it will be treated in this workshop, is thus a decision-oriented activity. Such an orientation leads to the following definition of evaluation, a definition which applies to each of the four stages of the Center's framework:

> Evaluation is the process of determining the kinds of decisions that need to be made about an educational enterprise; selecting, collecting, and analyzing data to inform the decisions; and reporting results of the evaluation to appropriate audiences.

One of the first tasks facing an evaluator is, then, to identify the kinds of decisions that the evaluation is to serve. Only after this is accomplished can evaluators choose and implement an appropriate set of evaluation procedures. The Center's evaluation framework is intended to help evaluators pinpoint the classes of decisions that are to be made; it also should help to guide the selection of an evaluative approach that will yield the needed information. Four major classes of decisions, and the evaluation activities appropriate for each, are pre-

sented in the following overview of the CSE evaluation framework.

CSE EVALUATION FRAMEWORK

Decision to be made	Appropriate Evaluation Activity	Activity Description
Identify "high priority" goals that will serve as the basis for new or revised educational programs	NEEDS ASSESSMENT	Select from among a full range of alternatives the goals to which resources should be committed. Determine extent to which current program is achieving the "high priority" goals.
Decide the means for achieving previously selected goals.	PROGRAM PLANNING	Select and/or develop the program most likely to attain goals. Develop a plan, including evaluation procedures, for the proposed program.
Decide where improvements in an on-going educational program are needed.	FORMATIVE EVALUATION	Measure program processes and components to ensure program is following the plan. Assess achievement of program goals. Recommend improvements in program where needed.
Determine the overall effectiveness of a fully-developed program to aid decisions about continuing, expanding, or dropping program.	SUMMATIVE EVALUATION	Measure program outcomes and document program processes to permit informed decisions about overall effectiveness.

The overview just presented suggests that there is an orderly and logical sequence for evaluation activities, beginning with Needs Assessment, proceeding through Program Planning and Formative Evaluation, and concluding with Summative Evaluation. This logic does, in theory, hold true. In actual practice, though, an evaluator rarely has an opportunity to participate in all four evaluative stages for a single program. It is far more likely that services of an evaluator (or several evaluators) will be called for. It also is rare for the evaluator's mission to be explicitly defined, such as: "We have now completed a needs assessment; will you please assist us in program planning?"

The ambiguity inherent in many evaluation projects is a powerful reason for evaluators to have a good conceptual understanding of the evaluative activities associated with different phases of educational program development. The time line below relates typical concerns and activities of program development to the services that evaluators might be called upon to provide. Note that the diagram does not indicate an exact proportion of time that should be devoted to each phase of activity. Clearly, for example, the time frame for evaluating a two week instructional program will be considerably condensed compared to a semester-long program.

LIFE OF A PROGRAM AND RELATED EVALUATION ACTIVITIES

	NEEDS ASSESSMENT	PROGRAM PLANNING	FORMATIVE EVALUATION	SUMMATIVE EVALUATION
PHASE OF PROGRAM DEVELOPMENT	·goals/needs that should be addressed are to be identified ·a specific program has not yet been chosen	·high priority goals have been identified ·a method for achieving goals must be selected ·the evaluation of the selected method must be planned	·a program has been selected and put into operation ·program is still "fluid;" modifications will be made	·program has been running for awhile; procedures and methods are fairly stable (no major changes will be made)
WHAT THE EVALUATOR CAN DO	·help to clarify goals ·assist with identifying needs ·measure effectiveness of current programs vis-a-vis important goals	·facilitate planning decisions ·help program planners develop, adapt or select program to meet high priority goals ·build the evaluation component into the program plan	·determine extent to which program is running as planned ·measure program's progress in attaining its goals ·provide recommendations for improvement	·document program processes ·measure attainment of specified goals ·determine whether program is promoting any unplanned effects ·compare to competing program(s)
EVALUATION FINDINGS NORMALLY REPORTED TO	·people responsible for allocating instructional resources (e.g., funding agency; district, state, local administrators) ·people responsible for instructional planning	·people responsible for instructional planning (e.g., teachers, curriculum specialists, administrations) ·community	·program director ·program staff	·people responsible for allocating instructional resources ·groups concerned with program effectiveness
EVALUATION FINDINGS TYPICALLY RELEASED	·when the evaluation is completed	·evaluation plan within program plan submitted to funding agency	·throughout course of the evaluation	·when the evaluation is complete

A-5

Section A

Exercise 1

Categorizing Evaluation Activities

Directions

The Center for the Study of Evaluation has developed a four-stage framework model of educational evaluation. The four stages are:

1. Needs Assessment (NA)
2. Program Planning (PP)
3. Formative Evaluation (FE)
4. Summative Evaluation (SE)

During these stages, the individual(s) responsible for conducting an evaluation focuses on certain activities. In this exercise, you will be given a list of 10 specific evaluation activities. Your task is to select the evaluation stage in which each activity belongs.

Place the initials of the appropriate evaluation stage (NA, PP, FE or SE) in the blank before each activity.

Section A

Exercise 1

Categorizing Evaluation Activities

_____ 1. Ms. Eve Alewaiter is in the process of developing ways to identify important objectives.

_____ 2. Opening the packet containing the evaluation results, Ms. Sue Peervisor was overjoyed that she could finally make the decision to continue this program or go back to the old program.

_____ 3. Ms. Eve Alewaiter was busy developing procedures to identify which instructional program might be used to meet the district's math objectives.

_____ 4. Ima T. Cher spent some time during the program giving unit post-tests to determine how well en-route objectives were being met. These results were then shown to Eve Alewaiter.

_____ 5. The program planners were provided with interim reports from Ms. Alewaiter to let them know how the program was functioning.

_____ 6. In the initial stages of the evaluation, Ms. Eve Alewaiter spent some time deciding on an appropriate evaluation design to use to determine the effectiveness of the program.

_____ 7. Eva was delighted to prepare a summary report in which students in the new program appeared to achieve the outcomes better than students in the traditional program.

_____ 8. A very important task undertaken by the program staff was to identify the procedures they would use to determine the priorities of important objectives.

_____ 9. Before the program began, Ms. Alewaiter spent some time with a statistician colleague discussing how the data from the program might be analyzed and reported.

_____ 10. Teachers, students, and parents were given questionnaires which asked them to assess the worth of the current school programs.

SECTION 8

INTRODUCTION TO THE SIMULATION

SECTION B

Introduction to the Simulation

Evaluation Workshop III proposes to clarify evaluation concepts and techniques by providing opportunities for you to discuss and apply them. Area E, a hypothetical administrative unit of the New Hampton Unified School District will be the workshop simulation. During the workshop, you will be working individually and in teams to formulate a new experience-based Business Education Program for the district and to determine the scope and context of the program's evaluation. The exercises in the workshop approximate the steps that the staff evaluator, Dr. Peter Slate, follows in his evaluation. You and Dr. Slate will examine the program planning at J.F. Kennedy High School from the evaluator's perspective. Thus, when you have completed the workshop, you will have a "case study" of the evaluation.

The following material will provide you with the background information necessary to your participation in this simulation.

In this evaluation, you will encounter six prominent characters: the evaluator, the area superintendent, the principal and three teachers.

The evaluator, Dr. Peter Slate, has just joined the District's Research and Evaluation Division. Because of his excellent reputation, the Area E superintendent asked for Dr. Slate, in particular, to help plan and evaluate the new Business Education Program. Dr. Slate has an enthusiastic and open approach to evaluation. This attitude has made it easy and enjoyable for people to work with him in the past.

Dr. Samuel I. Taft is the Superintendent of Area E and also the coordinator of the Business Education Program Planning Committee. He has been both a teacher and a principal in an inner city school and thus has some sensitivity to the particular concerns that his Area E teachers and principals encounter. He is an innovative man and this Business Education Program constitutes his first major project as superintendent. Since Dr. Taft is convinced of the excellence of the experience based concept, he has high hopes for it. Consequently, his attitude toward the evaluation is rather reserved, as he suspects that a negative report might affect the longevity of the project.

Florence Williams has been the principal of J. F. Kennedy Senior High School for seven years and has been a principal for a total of twelve years. Despite her quiet personality, Ms. Williams has an intense concern for inner city students. She is against any new Business Education Program because she feels that it is more important to devote time and energy towards equalizing the level and quality of inner city education with that of other areas. Although she recognizes the need for an experience based business education, she maintains that its essential elements can be just as easily conveyed in a traditional business education class. She encourages the evaluation in hopes that it will document enough flaws to warrant the discontinuation of the program. This outcome would permit efforts to be channeled toward improving J.F. Kennedy's basic programs.

Marlene Blakely has taught business education at J. F. Kennedy for ten years. She has a buoyant personality and has a tendency toward leadership which makes her the obvious leader of the three member planning team. She has participated in a summer business education workshop and has a positive

attitude toward the current program. She has a guarded opinion of the evaluation, however, as she has had no previous experience with program evaluation.

Joe Crennin is a first year business education teacher at J. F. Kenredy. He too attended the summer business education workshop. He has an enthusiastic and determined approach to his work. Mr. Crennin is a young teacher and doesn't have tenure at the school. He is rather concerned about the effect that a new and unproven program will have upon his chances for tenure, especially since not as many teachers may be needed for the new program. He welcomes the evaluation and hopes that it will show that the new program is not suited to Area E.

Martin Smith has taught business education and economics at J. F. Kennedy for five years. Of the three planning team members, Mr. Smith has had the most practical experience with business education. He works as a volunteer with the Better Business Agency, a citizens' advisory group. As a result of this participation, Mr. Smith has numerous and valuable contacts in the business field. During his five years at Kennedy, he has been developing and assembling business materials for his business education classes. Mr. Smith has positive feelings toward a new business education program but feels that evaluation as a whole is a waste of time.

As we have noted, Area E is an administrative unit of the New Hampton Unified School District, a large metropolitan school district like New York City, Chicago or Los Angeles. Located within the "inner city," this

unit includes twenty-nine elementary schools, eight junior high schools, one intermediate school, and four senior high schools. The administrative head of each Area, known as the Area Superintendent, controls the allocation of available discretionary funds for special programs.

Dr. Taft has decided to allocate money for the development of a high school business education program. The results of a needs assessment, which involved the principals, the school officials, parents, teachers, and students of the four high school prompted his decision. All four schools completely agreed on the top-rated objective —— "to develop skills in reading, writing, speech, and listening" —— and closely agreed on all other goals. Furthermore, each of the four schools emphasized the need to teach young people practical business operations and to help them become capable managers of their money, and resources. At this time, Area E's courses in business education and economics tend to reach relatively few students and the courses are too theoretical and abstract. Thus, the students who graduate from these high schools lack basic skills in obtaining and using information that would enable them to establish themselves as responsible business people in the community. In response to the results gathered by the needs assessment, Dr. Taft began to organize the systematic planning of a business education program for the four high schools in Area E.

Personnel Summary Chart

Name	Position	Attitude toward Program	Attitude toward the Evaluation
Peter Slate	New District Evaluator	Open	open and enthusiastic
Samuel Taft	1) Superintendent, Area E 2) Coordinator of Planning Committee	positive	reserved (negative report will affect longevity of program)
Florence Williams	Principal, Kennedy High	negative	positive (hopes report documents program flaws for subsequent discontinuation of program)
Marlene Blakely	Business Education Teacher	positive	guarded (no experience with evaluation)
Joe Crennin	Business Education Teacher	negative (can eliminate tenure opportunities)	positive (hopes it will discredit program)
Martin Smith	Business Education and Economics Teacher Better Business Agency Volunteer	positive	negative (evaluation a waste of time)

SECTION C

OVERVIEW OF PROGRAM PLANNING

SECTION C

Overview of Program Planning

C. Overview of Program Planning
   1. Outline of Program Plan
   2. Program Planning Staff
   3. Role of the Evaluator
   4. Exercise 2

At the completion of the Needs Assessment stage of an evaluation, the desired goals and objectives of a program have been clearly outlined. The next step then, is to plan a way to meet these objectives. This brings us to the program planning stage. During program planning, a planning staff must complete two major tasks: 1) either create, adapt or select a program to achieve agreed upon goals and 2) build in the evaluation of the program. Although this workshop discusses program planning in the educational setting, a program plan can be written for any kind of program, e.g., an energy conservation program or a staff reorganization program.

Working as a team, the evaluator and planning staff should produce a written program plan that address both major tasks by the end of this stage of evaluation. The components of a program plan are outlined below:

Outline of a Program Plan

I. Summary

   A. Rationale, purposes, justification for the program, brief description of the general goals, procedures, and evaluation.

II. Program Description

   A. Instructional Characteristics

      a. Program goals

         (1) general goals and specific objectives

b. Major program activities

(1) description of activities

(2) sequence of activities

(3) instructional components of each major activity

B. Administrative Characteristics

(1) pupils served by program

(2) resources needed - materials, facilities, equipment, personnel

(3) schedule - when do the program's activities take place, duration of activities

(4) job responsibilities of those implementing the program

III. Evaluation Plan

A. Purpose

(1) general formative or summative information needed by decision makers

(2) specific evaluation questions - outcome, cost, implementation, research questions

B. Design

(1) participants - who will be measured

(2) measures - the tests, questionnaires, etc. to answer each evaluation question

(3) data collection plan - when measures will be gathered for each evaluation question

C. Budget for the evaluation

D. Data Analysis

E. Reporting Evaluation Results

Program Planning Staff

In order to write the instructional, administrative and evaluation components of the plan, the planning staff usually includes people with a variety of skills.

Some of the planners are usually instructional experts like teachers and curriculum coordinators who, because of their special skills, will mainly concern themselves with instructional goals, descriptions of teacher behavior, students behavior and materials. If the program staff has decided to create a program, their instructional knowledge will be especially important. If the planning staff decides to select a program, their major concern will be to find out what programs are available that focus on high priority goals.

Administrators are also often included as program planners. They will usually be concerned with directing the planning activities -- setting and keeping to a planning budget, coordinating planning logistics and informing the community about the planning effort. Administrators will also be concerned with finding a program that conforms to the districts' cost and logistical constraints.

Role of the Evaluator

The evaluator assists in both components of program planning. As previously mentioned, these components are: (1) to develop, adapt or choose a program to meet specific educational needs, and (2) to plan the implementation and evaluation of this program. Program planning evaluators may be either internal regular staff or external independent consultants. A major difference between internal and external evaluators is often the point at which they become involved in the evaluation. The internal evaluator may be on the job sooner and longer; the external evaluator may join the

evaluation effort later in the project and not have a long term involvement. Frequently the program planning evaluator will be an internal evaluator who is a regular staff member. The internal evaluator may have participated in the needs assessment and has a long term working relationship with local people, enabling him to be of help in developing (or selecting) and implementing a program. Ideally, the evaluator, whether internal or external, should be involved in the early stages of evaluation.

The evaluator must have a variety of skills for the many tasks required in an evaluation. Note that the tasks for program planning vary depending on whether the program planners are developing or choosing a program. Though the task may be handled by the planning staff, the evaluator, or both, the evaluator does not have to be knowledgeable in all tasks to be of assistance. Instead, the evaluator may serve as a guide in the selection of appropriate program objectives and other program considerations which would help facilitate the evaluation process.

Common Evaluator Tasks in Program Planning

Program Development

__ help planners clarify their objectives

__ examine relationship between objectives and activities

__ examine relationship between measures and the objectives and activities

__ estimate time, money and resources for program development

Program Selection

__ lay out existing programs in such a way that information about them can be prepared

__ help planners list relevant considerations which they must weigh appropriately to select a program

__ suggest techniques to use in choosing programs

__ analyze programs from stand point of their development and evaluation information

__ determine adequacy of testing, diagnostic and record keeping instruments included in programs

__ help planners allocate time, money and resources for program selection

Program Evaluation

__ help planners allocate time, money and resources for the upcoming evaluation

__ suggest, clarify and agree upon the evaluation questions (e.g., outcome, implementation, research, cost and unanticipated outcome questions)

__ examine relationship between measures and the objectives and activities

__ specify information collection procedures (e.g., interviews, questionnaires, observations, cognitive and affective measures)

__ decide on design and schedule to administer data collection procedures

__ determine the sample to be used

__ decide on data analysis

__ determine schedule for reporting analysis

Section C

Exercise 2

Tasks and Responsibilities in Program Planning

The chart on the previous page outlined the usual tasks in program and evaluation planning. In this exercise you will assign each of these tasks to either the planning staff, the evaluator, or both.

The workshop leader has asked you to take the role of either the planning staff or the evaluator Peter Slate. Refer to the Introduction to the Simulation (Section B) and the chart (p. B-5) for the background and attitudes of the personality you have assumed. Each task should be assigned from the perspective of the person you are role playing. In the appropriate space next to the job description, write an E if you think the job is the responsibility of the evaluator, a P if you think it is the responsibility of the planning staff, and PE if you think that the job should be shared by both planners and evaluator. Remember to keep in mind the attitude of your character toward the program and evaluation.

Note that the Job Descriptions listed are not comprehensive but do indicate main activities. Other considerations do exists as, for example, how many copies of the report are to be made or who does the duplicating of material.

Section C

Exercise 2

Jobs and Responsibilities in Program Planning

Job Descriptions	Person(s) Responsible
Decide on objectives, instructional materials and teaching strategies, student activities	
Ensure that objectives are clear, so they can be measured	
Define program resources (time, money, personnel); constraints and facilitating factors (new equipment); special requirements (legal questions).	
Determine the extent of evaluator's involvement in program planning	
Develop or select tools and procedures for selecting among existing alternative programs	
Choose a program to meet the District's needs	
Place restrictions on the running of the program (e.g., random assignment of students)	
Specify the purpose of the evaluation	
Schedule the evaluation	
Describe what evidence will be accepted to judge the program's effectiveness	
Select methods and measures for assessing program success	
Identify data-collection requirements	
Plan the data analysis	
Write the program plan	

SECTION D

PROGRAM PLANNING AND EVALUATION CONTEXT

Section D

Program Planning and Evaluation Context

D. Program Planning and Evaluation Context
1. Context Question
2. Exercise 3
3. Document File

One of Peter Slate's first tasks is to familiarize himself with the context or environment in which program planning, takes place. In determining the program planning context evaluators must determine what kinds of resources are available for the program and its evaluation.

To determine the nature of a particular context, evaluators should ask relevant questions that will be important for program planning and evaluation. Evaluators should discuss these questions with the program staff and the decision makers before planning begins. This will help increase the usefulness of the evaluation information for program planners and decision makers by helping the evaluator come up with the "best fit" between the program staff's and decision maker's goals and good evaluation practice.

The evaluator should be prepared to discuss the following with the program's staff and decision makers

1. program characteristics
2. program evaluation requirements
3. school organizational structure
4. community organizational structure.

In each of these areas, the evaluator should raise a series of questions aimed at determining if everyone is clear and agrees upon such issues as the goals of the evaluation, who will be responsible for what, procedures will be followed, who reports to whom. In some instances the line of

responsibility and procedural guidelines will be clear, in others the locus of responsibility or procedures to be followed might need to be negotiated depending on the particular program being evaluated.

Outlined below are some "start-up" questions that evaluators might try to answer when they first become involved in an evaluation project. Please note that the list is suggestive, but not inclusive, of all concerns that may be important in different settings.

Context Questions

I. Program Characteristics - Overall Description of the program.

A. Will a program be selected, developed or adapted?

B. What is the subject area?

C. What is the origin of the program?

1. How did it get started?

2. Who are the prime developers?

3. What was the major impetus for the program?

D. What are the general goals? How were they established? (needs assessment, mandated, other)

E. What are the specific objectives? How were they selected? (research, theory, rationale)

F. What kinds of students will the program serve?

1. For what age, grade level, and ability level is the program appropriate?

2. How will students be selected for the program?

3. What criteria, if any, determine their continuing participation?

G. What strategies (teaching style or instructional methods) will be included in the program to bring about program goals?

H. Who will operate the program?

1. How many teachers, administrators, clerks, aides, volunteers, and others will be running the program?

2. Do they have special training?

I. What is the allotted budget for program planning?

II. Program Planning and Evaluation Requirements.

A. How much time, money and personnel are available to plan and implement the program?

1. Does the budget make provisions for these activities?

2. Is it practical to plan and implement the program?

B. Is the necessary staff available to plan and develop the program?

1. If not, can they be employed?

2. Will consulting help be needed?

C. What is the general purpose of the evaluation?

1. What decisions about the program will be made based on the evaluation?

2. What is the place of evaluation in the organization? Is it conducted at the local, district or state level?

3. What is the general scope of the evaluation (e.g., number of schools, number of students, number of programs)?

4. Is there a hidden agenda?

D. How much time is allotted to conduct the evaluation?

1. What are the reporting deadlines?

2. How much time is available for data collection?

E. Are there mandated evaluation requirements to comply with funding agency guidelines and/or laws?

F. What kind of evidence of program success is either requested, desired or possible?

1. What kinds of potentially useful data can be obtained? (interviews, on-site observations, special performance tests, questionnaires, etc.)

G. What audiences are there for information about the program and the evaluation?

1. What are the audiences for interim program information and the final evaluation report?

2. How frequently should reports be made?

3. How many copies of the report need to be made?

III. School Organization and Structure

A. Who are the main decision makers connected with program planning and evaluation?

1. Why are they involved in program planning and evaluation? (self initiated or required)

2. What are their motivations? (e.g. to determine if a program is successful vs. showing that a program is successful)

3. What is their knowledge of evaluation?

4. What is their committment to/attitude toward program planning and evaluation?

5. Do they agree with the prepared requirements of the evaluation?

6. What is their managerial style (shared, authoritarian, timid, aggressive)?

B. Who does the evaluator work with and report to?

C. What are the formal and informal lines of communication within the program?

1. Who are the opinion leaders in the program?

2. To whom do formal/informal reports go?

D. Are the individuals who will participate in the planning and implementation of the program willing to first systematically plan and then change the program if evaluation results show that changes are needed?

E. What is the status of the program's planning staff?

1. How long has the group been in existence?

2. How strong is their committment to program planning/evaluation?

3. How strong is their evaluation expertise?

4. How strong is their support in the community?

E. Are there other special interest groups in the school structure with involvement in and attitudes toward the program and its evaluation?

IV. Community Organizational Structure

A. What involvement with program planning do parents, business people and other community groups have?

1. What information about the planning effort and the evaluation does the community have?

2. What are the community's attitudes toward the program and the evaluation?

B. Who are the potential critics of the program and/or the evaluation?

1. What are their concerns?

C. What "political" conditions in the community, district or state exist that may influence program and evaluation activities (e.g., Is a changing school enrollment creating needs for expanding or deleting programs? Is a local newspaper editorializing for local school reforms?)

Section D
Exercise 3

Program Planning and Evaluation Context

Directions: As you read Documents 1-8 look for the information provided about the scope and context of the evaluation. Fill in only the information indicated on the answer sheet that corresponds to your team identification (A, B or C). Write the number of the document containing the information in the right hand column under the heading Document Source.

TEAM MEMBER A

I. Program Characteristics	Document Source
What is the subject area?	
_____	_____
What are the general goals?	
_____	_____
_____	
_____	
How were the goals established? (Needs assessment, mandated, other)	
_____	_____
Is there a program selected already?	
_____	_____
Is it in operation?	
_____	_____
Will one be selected or adopted?	
_____	_____

What kinds of and how many students does it serve? [SES (socio-economic status), age, ability, ethnicity, number]

What is the teaching method, style of the program?

IV. Community Organizational Structure

What involvement with the program do parents, businesspeople and other community groups have?

What information about the program and evaluation do they have?

What attitudes toward them do community people have?

TEAM MEMBER B

II. Program Planning & Evaluation Requirements	Document Source
What is the general purpose of the evaluation?	
_____	_____
_____	
What decisions will be made?	
_____	_____
How much time and money are available to plan the program?	
_____	_____
How much time is allotted for the evaluation?	
_____	_____
Are there mandated requirements for the evaluation?	
_____	_____
_____	
_____	
What kind of evidence is required, desired, and possible?	_____
_____	
_____	
_____	
_____	

TEAM MEMBER C

III. School Organizational Structure — Document Source

Who are the decision makers?

_____ _____

What are their positions in the organization?

_____ _____

Why are they involved in the evaluation? (self-initiated, required)

_____ _____

What are their motives for having the program evaluated?

_____ _____

What is their knowledge of evaluation?

_____ _____

What is their committment to/attitude toward evaluation?

_____ _____

Do they agree with the proposed requirements for the evaluation?

_____ _____

How much teaching experience does the planning group have in business ed?

_____ _____

How much experience does the group have in evaluation?

_____ _____

What attitudes do the team members have toward evaluation?	Document Source
What is the place of the evaluation in the organization? (local, district, or state level)	
Who does the evaluator report to and work with?	

Section D

Document File

New Hampton Unified School District

Memorandum

TO: Area E. Senior High School Principals

FROM: Samuel I. Taft, Superintendent of Area E

SUBJECT: Status Report on the Business Education Program

As a result of the needs assessment conducted in your school, Area E has organized a planning committee to develop a one semester experience-based Business Education program for our senior high school students. The general goal of the program is:

To provide students with information about the responsibilities of both a business person and a consumer in a business interaction.

As principals of the inner city schools that comprise Area E, I am sure that you have realized the very particular and difficult problems that inner city residents face in entering the world of business. The planning committee has considered a number of documents. The committee also felt that you might find the information contained in the various documents we have used helpful to your staff and your School Advisory Council in discussions of the Business Education Program. Copies may be obtained by telphoning my secretary, Mrs. Trittle at 839-1511, ext. 1.

At the first meeting of the planning committee on Friday, June 5, 19XX we agreed that in order to ensure the success of our planning efforts, it will be necessary to add an evaluation service to our committee as soon as possible. After interviewing several candidates, we agreed that Dr. Peter Slate, of the District's Division of Research and Evaluation, will head the evaluation staff.

SIT/lt

Conversation between Dr. P. Slate, S. Taft, and F. Williams
Monday 10 a.m.

Samuel Taft: Good morning, Dr. Slate...

Peter Slate: Pete, please.

ST: Okay, Pete, I'm glad you could make it. You remember Florence Williams, the principal of John F. Kennedy High School, from your interview, don't you?

PS: Sure, glad to see you.

ST: Anyway Pete, before we get into an actual discussion, let me outline a few of the practical considerations for you.

PS: Good idea. That'll help.

ST: All right. As you know we need to develop or select a good Business Education Program, particularly for inner city schools like these in Area E. Based in last year's needs assessment, we know that high school graduate have little knowledge of business practices. Specifically, we'd like to try out two Business Education classes each semester... We have $XX for program planning and $X for the evaluation. Um, let's see... I guess that's it. Now, do you have any questions for us?

PS: Yeah, a few. For one thing, how much time do we have for planning and for the evaluation?

ST: Three months for the planning; one semester for the evaluation.

PS: Is the evaluation mandated?

ST: No.

PS. Who is planning the program?

F. Williams: We have three teachers involved in the planning. Two of these, Marlene Blakeley, who's been with us for 10 years, and Joe Crennin, who's been here for three years, have attended a Business Education workshop. The third, Martin Smith, who's been teaching economics and business education here for five years, is a volunteer in a citizen's advisory group--at the Chamber of Commerce.

PS: Well, they seem well qualified. Are they as enthusiastic about the idea of an experience-based business education program as they sound?

FW: Marlene and Martin like the experience-based concept, but Joe and I think it's a waste of time.

PS: Okay, just exactly what did you have in mind as the purpose of the evaluation?

ST: Since this is a new program, we want to collect information about how the program is being put into operation. We want to see how it's going in order to change or improve the program. Eventually, I would like to extend the program to all of the district high schools.

PS: So, you want lots of feedback as the program progresses.

ST: Yes, and it has to get to the program staff quickly enough so that they can make changes in the program. Of course, Florence and I'll want to see it too.

PS: Sure. Another thing, does any one else get the information?

ST: No. I don't want any interference when we're just beginning. We plan to work with business people, but we don't want any adverse information to affect their judgment of the program.

PS: How to the business people feel?

ST: Most are for it. Some have questions.

PS: Who will I be working with?

FW: Directly with me and the planning teachers.

PS: And, this information, how frequently will you expect it?

ST: Well, we'd like it monthly. We'd also like a report at the beginning and the end of the programs; the final report has to be in by June 30th because we have a meeting with the School Board at that time. We'll want to be able to discuss your report with them.

PS: Who'll see that final report?

ST: Just Florence, the staff and I.

PS: Well, what about all the others? The business people, the parents, the board and the other teachers?

ST: That all depends on what the results look like.

FW: We'll probably have to write something to them.

ST: I guess so, but we'll think about that later.

PS: What kind of evidence does the Board want?

ST: Student performance on a program-based criterion-referenced test and reports of success from business people.

PS: Well, for the moment that just about wraps it up. I'd like to talk with the planning team, then we may have to meet again.

ST: All right then. Florence, can you think of anything else? No, okay then. Pete, it was nice talking with you and we'll be getting in touch later.

PS: Right. Thanks for answering all my questions. It was very helpful. 'Bye Florence. 'Bye Sam.

FW: Goodbye, Pete. See you later.

ST: 'Bye.

Telephone Conversation between Marlene Blakely and Dr. Peter Slate

Peter Slate: Hello, Ms. Blakely.

Marlene Blakely: Yes.

PS: This is Pete Slate, the evaluator Mr. Taft hired to help with your Business Education Program. I'll be working with you soon.

MB: Oh, yes. Florence, our principal, told me about you. What can I do for you?

PS: Well, I have some questions that I hope you won't mind answering.

MB: Sure, go right ahead.

PS: Great. For one thing, I was curious about Sam Taft and what he really feels about the program.

MB: Well, he started out as a teacher and went to a principal and then became a superintendent. So, he's familiar with all our positions. As for the program, he's definitely in favor of Business Education.

PS: I see. I was just wondering because when I met with him the other day, he seemed a little restrained.

MB: Oh, I don't think that was the program, it was probably a reaction to you.

PS: What do you mean?

MB: Oh, not you personally, just evaluators in general. It's just that he's afraid that any negative findings might influence the committment of the community and the teachers to the program. He's afraid that an evaluator may point out weaknesses in the program.

PS: Well, I'm going to . . . but judiciously, I'm not a hatchet man! I'm as eager as you are to see such a program materialize.

MB: I don't want you to get the wrong idea of Mr. Taft. He's a bit leery of evaluators because of his previous experience with evaluation.

PS: Oh, could you tell me about that?

MB: Well, his only contact with evaluation has been during the state achievement tests, which unfortunately revealed the low inner city scores. This knowledge, or rather the lack of it, angered the parents, who blamed Mr. Taft. They feel that it is his responsibility to improve their children's scores.

PS: I see . . . I can appreciate his reticence. Well, I'll be meeting with him again early next week and I'll try to put his fears to rest. I hope he'll see that I'm here to help the program, not to "leak information to the outside," although new programs are bound to incur some negative reactions, whether I publicize the results or not. I hope he and I will be able to work together. As you know, his opinion carries a lot of weight and I'd like his attitude toward the evaluation to positively influence the others involved with the program.

MB: Oh, I don't think you should worry about that. He does have a strong personality, but I think the meeting with you is indicative of his style. He's very responsive to our opinions.

MB: For instance, he conferred with the principals and teachers about the need for Business Education and about what program to select.

PS: Well, all right. Speaking of teachers, I know that you and some of your planning team partners are very enthusiastic about the program and that you're well qualified to develop it. Let's see, it was you and Mr. Crennin who were involved in the workshop, right?

MB: Yes. That's right.

PS: I'm curious to know how you feel about the evaluation aspect.

MB: Well, I think I can speak for Martin, when I say that we are well aware of the need for evaluation, but that we share Mr. Taft's concern that a negative evaluation could hurt the program or interfere with its operation. This will be our first encounter with professional evaluation, you know.

PS: I certainly understand your feelings and I can only stress that I'm here to help the program, not to kill it. I'm sure between all of us that we can develop a good program.

MB: Oh, one other thing. Joe expressed a concern that this evaluation will reflect upon our teaching ability.

PS: Don't worry about that. After all, this is a new program all around and no one is expected to perform superlatively the first time out.

MB: Oh, good. I see.

PS: Did you have any other concerns?

MB: No, not really.

PS: Well then, I think I've taken up enough of your time. Thanks a lot for all your help. I look forward to seeing you and actually getting the program going.

MB: Me too. I was glad to help out.

PS: Okay, then, I'll be talking to you.

MB: Right. Goodbye.

PS: Bye.

TO: Joe Crennin

FROM: Florence Williams F.W.

The attached letter from Mr. Jones and Mr. Salazar may answer some of your questions about the program. It looks as if we will get a warm welcome from business people.

Flo -
I got the point. But you should talk to Bob McBride at City National. He has some strong reservations

Joe

Ms. Florence Williams
John F. Kennedy High School
New Hampton, State XXZIP

Dear Ms. Williams:

The decision of your school to engage in Business Education has come to our notice. We are two businessmen in your area who own small businesses and we would like to congratulate you on this wise decision. We cannot emphasize too strongly the importance of Business Education. To back up our feelings, we and six of our business associates would like to volunteer our help.

We propose taking on students as "interns" in our businesses so that they can observe, first hand, what goes into running a business. Our associates come from various businesses including: a bank, a furniture store, a credit company, two grocery stores and three department stores. The students will get to see the rights and responsibilities of both the consumer and business manager from the inside. Our offer will make your program "experience based" and more practical. If we can be of any further help, please feel free to call us.

Sincerely,

Robert "Lightfoot" Jones
Lightfoot Shoes

John Salazar
Salazar Drugs

RJ & JS/ic

Excerpts from Parents Advisory Council Report

. . . .voted to support the development of a Business Education program for the senior high schools. The program is to be characterized by its practical orientation and is to be "experience-based." That is, each student who participates in the program will be given the opportunity to spend one or more school days working in local businesses. . . .

. . . .To minimize the costs of the program, the use of films, filmstrips, etc. to teach basic business skills will be discouraged, except where the district or the Area has already purchased such devices. For example, there are several business math filmstrips currently in use in Area E. . . .

. . . .voted their preference for selecting a program that has been tried out, at least in part, with students in metropolitan areas. Although testimonials from teachers and students are considered important, more quantitative evidence of achievement of objectives must be provided.

Editorial from the Publication of the
Association of Teachers

BUSINESS EDUCATION: A Costly Reflection of Reality

Area E is involved in another "innovation": Business Education. As we all know, this subject may be best learned through application and not through theory or model situations. We are against Dr. Taft's intent to waste valuable educational time on material that is best outside the school.

We have several reasons for opposing a luxury such as Dr. Taft proposes. First, our schools still need to emphasize the basic skills, reading, writing, and mathematics. Students' expertise has been shown to be declining in these areas. Secondly, we contend that business education occurs naturally outside the classroom walls. Students, from the time they purchase their first ice cream bar from the ice cream truck, are exposed to business transactions. They are continuously instructed in business behavior outside of the educational institution. Anything that a Business Education program could teach the students about the pitfalls of business would remain purely theoretical until the student encountered it in real life. Moreover, nothing ensures that the student will make the necessary connection between the two. Thirdly, any cost incurred in the construction of a Business Education program could be more beneficially applied to the reinforcement of a basic skills program. We know that most of you feel the same way about the institution of a Business Education program. The following information should prove quite enlightening.

Dr. Taft distributed an "Attitude to Business Education Program Questionnaire" to all 400 teachers in New Hampton's High Schools. 240, or 60% responded. The Association sent out the same questionnaire to our members

(even though that meant that some of you probably did double work). Of the 205 members, 60%, or 123, members responded. The numbers look like this:

Statements	Association's Teachers			Taft's Teachers		
	Agree	Uncertain	Disagree	Agree	Uncertain	Disagree
1. Business Education programs should be instituted at some of the Area E schools	10	80	10	77	14	19
4. Business Education programs will lead to better business knowledge and practice	40	47	13	87	9	4

From the numbers it sure looks like Dr. Taft's teachers are in support of a Business Education program. But, why are the Association's teachers so uncertain? We sent out a follow-up questionnaire asking just that. We found that the majority of teachers who responded had different interpretations of what a few called the "hidden meaning" of the statements. For #1, for example, the teachers said that they thought no business education programs should be implemented in Area E schools, but were afraid to mark "disagree" since that might be interpreted as a "yes" to Business Education and that Business Education programs should be instituted at all district schools. The majority claimed that #4 looked good in print, but that they didn't really have evidence of its truth. By the way, twelve of the teachers who are in the area of the proposed Business Education Program are in our association. We did not actively search out their views because our questionnaires were anonymous.

We participated in last year's needs assessment, and we feel, for the preceding reasons, that, at this juncture, Business Education is a needless luxury. Dr. Taft should realize that, with academic necessities pressuring our teachers, life is a better instructor when it comes to business than our teachers can currently afford to be.

Document 8

July, 19XX

TO: Samuel I. Taft
Superintendent, Area E
School District

FROM: Peter Slate
Evaluator

SUBJECT: Preliminary Letter of Understanding

The purpose of this letter is to provide you with a report of our planning progress to date and to outline my understanding of my evaluation responsibilities to you and your program staff. When the plan for the actual evaluation is specified, I will again furnish you with a progress report.

Our general agreement and plans at this point consist of the following: The program, which will be put in operation in two classes each semester at John F. Kennedy High School, will consist of an experience-based Business Education program. The program, which will begin in September, will be subject to a formative evaluation during the year. While the program is in operation, I will be responsible for furnishing monthly reports to you and to the program staff which will describe my findings about the program, and will make recommendations, where necessary, for modifying and improving the program in terms of its implementation and progress toward student objectives. The evaluation will culminate in a final report that I will submit to you by June 30, 19XX.

Final decisions about the program to be selected, the measures to be administered, the schedule of testing and other information collection activities have not yet been determined. I will continue to meet with the program staff to help them decide how to make program selection decisions so that the best program to meet the area's needs will ultimately be selected. Once the program has been selected, I will submit a specific evaluation plan for your approval.

SECTION E

PLANNING THE PROGRAM

SECTION E

PLANNING THE PROGRAM

E. Planning the Program
   1. Introduction to Program Planning
   2. Clarifying Objectives
   3. Developing a Program
   4. Selecting a Program
   5. Program Selection Techniques
   6. Exercise 4

Introduction to Program Planning

As mentioned earlier, the evaluator has two primary tasks during program planning. This section of the workshop discusses the evaluator's first task -- to help program planners develop or select a program. The evaluator's second task, building the evaluation component into the program plan, is discussed in Section F.

Clarifying Objectives

Whether evaluators help program planners select a program or develop a program, one of their first tasks is assisting the program planners to

clarify the objectives the program should cover. Clearly stated instructional objectives serve several uses. They help program planners determine which objectives should be pursued and which objectives should not be pursued. They also help planners focus on the effects a program will have on learners. Clearly stated objectives make it easier to assess the effectiveness of a program and to evaluate whether or not program goals have been realized.

In working with program planners, the key to clarifying objectives is to help planners determine the observable behaviors, skills and competencies students should exhibit at the end of instruction.

In specifying objectives, it is useful to concentrate on objectives that describe a class of desired learner behaviors, rather than all specific behaviors. These expanded objectives, called content general objectives, are more efficient in program planning. For example, objectives might be:

1) To apply phonic rules in deciphering unfamiliar words

2) To distinguish between fact and opinion in persuasive writing.

The above objectives can be translated in two ways:

1) Objective - to apply phonic rules in deciphering unfamiliar words

a. Content general objective

The learner will be able to correctly read aloud, using the phonic rules, any 20 words randomly selected from a list of 220 unfamiliar basic words.

b. Test item equivalent objective

The learner will be able to correctly read alout 20 unfamiliar basic words beginning with the consonant cluster str.

2) Objective - to distinguish between fact and opinion in persuasive writing.

a. Content general objective

The learner will select statements of fact and opinion from

any piece of persuasive writing

b. Test-item equivalent objective - given 10 statements the learner will correctly identify statements of fact and statements of opinion.

The Case Study of Maria and Joe Glockenspiel

The following is a hypothetical real life example of applying the principles of content general objectives:

When Maria and Joe got married, everybody said that it was an ideal marriage. They had been sweethearts since they were in elementary school. Joe was an ambitious young man, clearly on the way to the top (wherever that is), and Maria's goal in life was to make Joe happy. For several years, life for them was all they had dreamed of. Joe had a good job, and Maria, following the precepts set by her mother, cooked, cleaned, and always managed to look pretty for Joe when he got up in the morning and when he came back from work at night.

Then Bill and Betty Sparta moved in next door, and the trouble started. Betty was an intelligent woman, and she and Maria hit it off right away. Betty was pursuing a degree in Indic Philology, and she fascinated Maria with her tales of university life, her friends, her professors, her interests and her successes. Maria got the idea that she, too, would like to go back to college. So she asked Joe. "Sure," was his reply. "Anything to make the little woman happy," was the way he phrased it to Betty's husband, Bill.

For awhile, Maria was happy in her dual role as housewife and student.

Then exam period came. She found she could no longer keep up the double pace, and, having absorbed some of the ideas of the liberated women on campus, she decided that she wanted Joe to assume more of the burden of keeping the house. She did not want him to do much - just enough so that she could get her studying done. At first, he and Maria fought about it; finally Bill and Betty overheard the fights, and decided to try to help. Bill got Joe to admit what the real problem was. He did not mind helping out, but he did not know how. His mother had always done all the work at home, and he did not even know how to boil water - let alone anything else. Maria enlisted the help of her professor in education, and the two of them came up with a set of behavioral objectives; things which Maria wanted to teach Joe so that he could help her about the house. This is what they looked like.

1. When the sink is full of dirty dishes, Joe will (a) fill the sink with hot water, (b) add one capful of detergent, (c) scrub each dish thoroughly with the dish sponge, (d) rinse the dish under running water, and (e) place the dish in the drainer.

2. Given the presence of lint on the rug, Joe will (a) remove the vacuum cleaner from the closet, (b) vacuum the rug using parallel strokes similar to the ones he uses in mowing the lawn, until the lint is completely removed.

3. Given an unmade bed, Joe will (a) draw the covers down to the foot of the bed, (b) smooth out the sheets tucking them in at the corners (c) draw each of the covers up to the head of the bed, smoothing them as he does so, (d) fluff the pillows (e) place them at the head of the bed, and (f) draw the bedcover over the whole.

4. Given a coffee pot, a can of coffee and a supply of water, Joe will (a) measure the coffee into the top of the pot, using one tablespoonful for each cup desired, (b) place as many cups of water as desired in the pot,

(c) plug the coffee pot in, and (d) wait patiently until the little red light goes on, indicating that the coffee is done.

5. . . . and so on.

Maria got tired of writing such specific objectives after awhile, and decided to ask her professor for help. "I'll spend my life writing these things," she wailed. "That," she said, "is because you are writing test items, not general objectives. You need more general statements of what you want." So Maria went home. This is what she came up with to replace the test items:

1. Joe will be able to clean up the kitchen after a meal.
2. Joe will be able to do general cleaning and dusting around the house when it is needed.
3. Joe will be able to tidy the bedroom in the morning.
4. Joe will be able to do the things which are necessary to prepare light snacks and meals.

Maria used the original test item equivalent objectives to test the general content objectives above. He passed with flying dishrags. Maria ultimately graduated with highest honors from the University. The subject of her doctoral dissertation was "An Instructional Model for Teaching Housekeeping Behaviors to an Ignorant Spouse." She then went on to teach Home Economics at Toluca U.

By now, however, there were deep philosophical differences between Joe and Maria, and they decided to separate. Joe, when last heard from, was using her dissertation on his new wife, whose mother had never allowed her to help around the house. He plans to live in the desert where they will have no neighbors.

THE END

Once the objectives are clearly stated, the evaluator should make the planners aware of the importance of keeping the desired objectives clearly in focus whether selecting or developing a program.

Developing a Program

In the simulation, the planning staff will be selecting one program from alternative programs. However, existing programs to meet a school's particular needs are not always available. Then the evaluator may help the school district develop or adapt a program. For example, before becoming a staff evaluator in Area E, Peter Slate helped the district develop an elementary consumer education program. This development process is summarized below:

Elementary Consumer Education in Area E

Helping the Planners Develop a Program

Peter Slate met with principals, primary grade teachers, and parents from two of the district's elementary schools. The parents and teachers were interested in teaching a consumer education unit in the lower grades. They felt that basic skills could be taught as part of the program and that children could learn concepts that would begin to provide a basis for combating the continual television advertising to which most young children are exposed.

The committee established some program objectives, and after reviewing a number of commercially prepared products, decided that the goals, costs, management requirements or instructional strategies were not consistent with the objectives they had defined. Peter Slate agreed to help the committee develop their own program. He emphasized the need

to create a prototype which could be tested on a small number of students. This strategy prevented some costly and time consuming errors.

Slate worked with the committee to assist them in developing instructional objectives to use as a basis for analyzing tasks. He then helped the committee to construct measurement instruments which would allow for the assessment of objectives attained. As the committee designed instruction based on clear communication of objectives, adequate practice, appropriate feedback and knowledge of results, Slate encouraged them to reference the instructional activities to the objectives and their assessment.

Selecting a Program

In contrast to the above example in which Peter Slate helped the progr planners develop a program, this simulation will follow Peter Slate as he assists the program planners develop systematic procedures to selecting a program. The program planners have researched existing business education programs and have found three that might be appropriate for Kennedy High's needs. To choose the best program, Peter Slate will suggest various comparison techniques which program planners can use to compare the different business education programs.

The following section describes some basic comparison techniques that Peter Slate might recommend. It also discusses the major advantages, uses (such as who will conduct the comparison) and limitations of each technique. The strategies discussed do not represent all those available and the descriptions of them are by no means intended to be a "Do It Yourself" guide. However, they will give you an overview of the range of techniques that evaluators have found feasible and sensible or many kinds of program analysis. It should be stressed that it is not the evaluator's responsibility to suggest program alternatives per se. His/her role is to help planners focus on their high priority goals and

to critically assess how well each potential program would meet these needs.

Program Selection Techniques

This section has been structured to cover a) who will conduct the comparisons b) what comparison techniques are available.
It is organized in the following manner:

A. Individuals Conducting Comparisons
   1. consultants/experts
   2. rating committees

B. Techniques for Comparing Alternative Programs
   1. recommendations
   2. Delphi technique
   3. program profile checklist
   4. systems approaches

A. Individuals Conducting Comparisons

1. Consultants/Experts

The primary advantage of hiring a consultant is acquiring expertise which the planning committee may lack in a specific area. Employing consultants may also prove to be a financial asset since it is generally more economical to hire high level expertise for a short time than to employ someone full time who lacks necessary qualifications. The consultant may merely substantiate the decisions of the planners. Even so, a second opinion from an outside source is valuable: the confidence gained from such support is a distinct asset. Furthermore, if the consultant does not encounter unanimity of opinion among the planners, s/he may be useful

in mediating the differences between committee factions.

Too often, however, the services provided by consultants are not used effectively, resulting in dissatisfaction with the consultant's performance. Employers can usually achieve a mutually satisfactory exchange of information if they follow the six guidelines below:

1. Specify in advance and in writing just what you want the consultant to do. Write the job description, including any restrictions, such as "it must be done by January 1" or "it must involve all 10 principals in the district."
2. Send descriptive information to the consultant about the school and the programs being considered before establishing a work relationship. On large projects, it is often better to pay the consultant for reading about the project before arriving so that the actual visiting time can be used profitably.
3. Require and pay for a written report from the consultant describing the analyses and rationale supporting any recommendation. These written reports will scare off many charlatans and provide you with a good record of what was said and why. Indicate the desired length and content of the report when you negotiate with the consultant.
4. Organize your schedule and the consultant's schedule so that you can meet in private <u>both</u> before and after the day of consultation to discuss the project.
5. Provide the consultant with a list of names and job titles of the persons on the planning staff. This procedure will produce an infor-

mality that often leads to a more productive atmosphere.

6. Find out if the consultant plans to use supporting help and who this will be. You may find that the consultant can not really do the job without assistance especially on jobs that take more than a day or two.

Rating Committees

A committee is another means of applying the selection criteria to competing alternatives. Like any tool, this approach can be misused. (The many jokes about committees include such examples as, "a camel is a horse put together by a committee.") When correctly used, however, the rating committee can be an effective method with advantages not found in other approaches.

The multi-person approach to selection ensures that a number of viewpoints will be considered in the rating process. In addition, people included in selection develop a feeling of direct involvement and have a greater acceptance of the final program. Care should be employed therefore, in selecting committee members. Often the committee may be some or all of the planning team. Groups that should be considered include district staff directly involved in the program (e.g., teachers, administrators, aides); persons with curriculum expertise in the area; persons with evaluation expertise; representatives from the community, especially groups with a particular interest in the program (e.g., parents with children in the program, civic groups that have been involved in the particular program, etc.); and students.

The major disadvantage of the committee is that it often proves an unwieldy tool: many person-hours can be expended in the selection of a program.

Some suggestions for improving the effectiveness of rating committees that do meet as a group are:

- Specify the decision-making role of the committee at the beginning (e.g., advisory capacity to Director of Curriculum).
- See that the committee uses a systematic approach to its assignment by encouraging them to state their objectives.
- Have the committee and the decision maker agree on a timetable for product selection.
- Recommend that a committee chairperson be chosen who has skills in group leadership. The leader should be able to: (1) prevent individual members from monopolizing group discussion, (2) help members stick to the area of individual expertise (a community representative may not be a curriculum expert), and (3) help the committee adhere to the agreed-upon time schedule.

B. Techniques for Comparing Alternative Programs

1. Recommendations
2. Delphi Techniques
3. Program Profile Checklist
4. Systems Approach

1. Recommendations

A frequently used technique for selecting among alternative programs is to solicit the recommendations of a subject matter expert. Whether the expert is a district or outside person, the value of the recommendation will depend on the basis on which the recommendation is made. For example, instead of simply asking "Which of these programs is best?" experts should at least be asked for written recommendations explaining the reasons that led to the recommendation. The major advantage of this approach is that it is easy to simply ask someone to give a recommendation. The disadvantages, on the other hand, are that the credentials of all the experts must be valid, the criteria that each uses must be explicit and meet the approval of the planning team, and, if several experts are involved, they may not agree. Therefore, unless there is a means of structuring the basis on which the expert recommends an alternative, this selection procedure is the least systematic, objective method available.

B. Techniques

1. Recommendations
2. Delphi Technique
3. Program Profile Checklist
4. Systems Approach

2. Delphi Technique

The term "delphi" derives from the oracle at Delphi, in ancient Greece, who forcasted its future. The term "Delphi Technique refers to the potential of making predictions based on polling opinions. The Delphi Technique, like the recommendations. enables the planning team to obtain and organize the views of a group of subject matter experts on the best program. However, unlike the recommendation procedures, the

Delphi approach offers a structured dialogue for predicting which program will be best. In fact, it is sufficiently structured so that the experts involved may never even meet one another, yet their ideas and expertise can be pooled.

The structured dialogue of this technique is presented in the form of a measurement instrument, either a questionnaire or checklist. It is therefore advantageous to discuss some of the considerations an evaluator should be aware of when helping the planners develop this instrument. For example, it is a good idea to format a questionnaire so that it is open to new and original suggestions for providing alternative ways of meeting the program goals. The questionnaire should encourage judgments about the efficacy of certain possible programs in terms of such factors as: instructional effectiveness, cost, feasibility and likelihood of success. The questionnaire should also include a section where the experts can specify the reason(s) for their judgments. This information could be very helpful for the planners as a future reference, especially if major areas of disagreement arise. Once the questionnaire or checklist has been developed, the basic steps for conducting a Delphi Study are as follows:

1. Distribute to a panel of experts the questionnaire or checklist about the various program alternatives that the planning staff have chosen. Include material describing the major aspects of each alternative.
2. Tally the results of the instrument to determine the areas of agreement and disagreement.
3. When major disagreement occurs, distribute a second questionnaire containing the various reasons given by the experts for their initial judgments. In this way, each respondent can comment on the opinions of the others.

4. Repeat the above procedure as often as necessary and feasible in an attempt to achieve closer agreement. Close agreement on a suitable program will hopefully result in a better prediction of program success.

The Delphi technique is especially useful when the situation is so complex and time so limited that the only economical and feasible way of gathering the necessary information is to trust the opinions of experts. The Delphi approach eliminates the necessity and expense of bringing together experts from different parts of the country, since it is more convenient to respond by mail than to meet at a prearranged time and place. The Delphi approach also avoids the biases and personality conflicts so often associated with face-to-face contact.

The major disadvantage of this approach, on the other hand, lies in the inevitable difficulties of interacting by mail -- forms forgotten, misplaced, lost in the mail. Further, there is no guarantee that the procedure will result in a consensus but simply provides a good climate in which people can "discuss" their views and change them if they think necessary, without the ego problems associated with winning or losing a face-to-face encounter.

B. Techniques

1. Recommendations
2. Delphi Technique
3. Program Profile Checklist
4. Systems Approaches

3. Program Profile Checklist

This comparison technique requires the evaluator and program planner to identify and agree on a set of questions or criteria that specify critical attributes which an extant program must have to be acceptable to the district's specific needs. While the specific criteria will vary according to the general nature of the program and the district, several categories of information may be valuable to include in any checklist. These include general information about the goals, cost, effectiveness and instructional philosophy of the program, the internal instructional quality of the program, and the managerial aspects of the program e.g., equipment availability.

A Program Profile Checklist consists therefore, of a set of considerations for comparing the merits of alternative programs or program components. It ensures that all key factors the planners and evaluators feel are important are considered when examining each alternative program. Its chief advantage therefore, is in focusing planners' attention on a common set of considerations for examining any alternative program.

The basic disadvantages of the technique are: failure to reach agreement among planners on the critical factors for selecting a program or the possible omission of important considerations. In addition, if two or more alternative programs have the features on the checklist, other means must be devised to choose among them.

In spite of these limitations, planners have found checklists quite useful in analyzing the merits of alternative programs. An example of a Program Profile Checklist is included on the following page.

PROGRAM PROFILE CHECKLIST

(Example)

A. General Characteristics

1. Goals - Do the program goals fit desired district goals?

(List goals)

2. Cost - Is the program practical to implement?
3. Effectiveness - Is there evidence that the program will increase student achievement?
4. Instructional Approach - Does the program use the style or philosophy of teaching desired?

B. Internal Quality

1. Instruction
   a. Are objectives stated measurably?
   b. Do materials present instruction on the objectives?
   c. Do materials and procedures provide sufficient practice and feedback for each objective?
   d. Do the materials present accurate, quality information on the subject area?
   e. Are the materials interesting?
   f. Are the materials easy to use?

2. Assessment

   a. Are frequent assessment measures provided?

   b. Are end of program measures available?

   c. Are record keeping devices easy to use?

C. Management

1. Are training and orientation procedures provided?
2. Are administrators, teacher, and students' procedures clearly defined?
3. Are there any chances that the program will interfere with other programs?
4. Will success of the program positively affect student performance on other programs?
5. Are the necessary personnel available to plan and run the program?
6. Is sufficient space available?
7. Are required facilities, equipment and materials feasible?
8. Is the program feasible to run in terms of logistics, communication, budgeting and time requirements?

B. Techniques

1. Recommendations
2. Delphi Techniques
3. Program Planning Checklist
4. Systems Approach

4. Systems Approaches

Several different kinds of procedures fall under the heading "systems approach." Since it would not be feasible to discuss them all let us consider only their common features.

The basic purpose of all systems approaches is to develop a systematic way of considering complex problems so that more enlightened objective, (and hopefully better) decision making can take place. The five basic elements in a systems approach are as follows:

1. Objectives must be identified.
2. Alternatives for achieving each objective must be specified.
3. Costs related to the achievement of these objectives must be specified.
4. A Model that displays the relationship between objectives, alternatives and costs must be constructed.
5. A Decision Rule must be developed for indicating from the application of the model, which would be the preferable alternative.

Let us assume, for example, that someone wishes to apply a systems approach to giving up smoking. First he would state the objective, which in this case is "to give up smoking by the end of the month." Next s/he

might list various alternatives for achieving this end, such as not buying cigarettes, taking pills, substituting chewing gum, etc. The third step would be to compute the costs associated with each of these alternatives, such as the price of gum for a month. The fourth step, constructing the model, is perhaps the key element since it combines and organizes the information in the three preceding steps. For example, a table can be made which shows the alternative procedures, and costs for each objective. Finally, the person must develop a decision rule or set of rules, such as "the best plan is one that allows me to stop smoking by the end of the month without spending more than $50 or gaining more than 10 pounds, or . . ."

To clarify the use of systems approaches in selecting among alternative programs, let us examine a systems approach which provides valuable data to the decision maker and is easy to use. This approach involves the construction of a decision matrix, and is described below.

There are several varieties of such matrices. However, they have a series of procedures in common for determining the relative merit of each alternative program on a numerical scale, such as 1 to 10. The obvious advantage of such matrices, therefore, is that they allow the planner to identify finally the relative overall value of each alternative. The major disadvantage of decision matrices, on the other hand, is the number of necessary (and often arbitrary) decision rules required to assign ratings of importance to each program characteristic. Although these matrices are somewhat more difficult to construct than the simple checklists described earlier, they do provide much more information about alternatives, and they are much easier to use than some of the highly sophisticated models and tools that might be employed. Thus, planners often select such matrices over other selection tools.

The basic steps in constructing a matrix are as follows:

1. List the objectives.
2. List and define all the general, instructional and managerial considerations used in determining the relative merit of each program alternative (e.g., cost, technical accuracy, teacher training).
3. Assign weights to each of the above considerations in order to specify the importance of each factor in relation to the others. A convenient weighting system is to distribute 100 points among the course considerations in terms of their relative importance.
4. Rate each alternative course on each of the item considerations, (e.g., 5 = very desirable, 1 = very undesirable).
5. Construct a decision matrix by multiplying the weight of each item (10-1) times the rating of each program on that particular item (5-1) to yield a score for each. For example, using Table 1 below, the score for goal agreement for the Ecology Course is 10 x 4 = 40, since 10 is the weight of importance assigned to goal agreement and 4 is the rating for how closely the goals of the Ecology Course mesh with district goals.

TABLE 1

WEIGHTS AND RATINGS
FOR COURSE CONSIDERATIONS
Ecology Program Alternatives

Course Considerations	Weight	Alternative Courses		
		Ecology Rating	Theory Rating	Traditional Rating
General Goals	10	4	3	5
Cost	17	5	3	4
Effectiveness	10	5	4	1

6. Arrive at a grand average rating for each program by summing each column of scores (not ratings) and dividing this sum by the total number of weights assigned. Thus, if the sum for the Ecology Course is 411, then one divides 411 by 100 (if 100 is the total points allocated) to get a final overall merit rating for the Ecology Course of 4.11. This number can be interpreted directly on the 5-point scale of desirability and/or compared to the final overall merit rating assigned to the other two courses.

Table 2, on page E-22, contains all the weights, ratings, scores, and final merit ratings for each of the three Ecology courses. This table indicates that two key factors in this approach are the weights assigned to each course consideration and the rating given to each course alternative on each consideration. Unfortunately, the subjective bases for determining these weights and ratings often are not as solid as they should be although the evaluator may help in firming them up with sophisticated rating and weighting procedures. Also, as in the case of the planning checklists, neglecting to include important considerations in the construction of decision matrices will seriously impair the validity of the results. The use of these matrices may require the development of additional techniques for handling special problems, such as exceptionally high or low ratings on some considerations, costs, or effectiveness. These are not insurmountable problems, and the evaluator can suggest sensible and effective ways for dealing with them. It should be remembered, however, that despite these difficulties, the information obtained from decision matrices is generally better than that gathered by simply asking the opinions of experts or using a planning checklist. Decision matrices are more valuable because the alternatives are fully evaluated and ordered in terms of their utility. When needed, more

TABLE 2

Final Rating of Alternative Ecology Programs

Program Consideration	Weight	Ecology		Theory		Traditional	
A. General Characteristics							
1. Goal Agreement	10	4	40	3	30	5	50
2. Cost	7	5	35	3	21	4	28
3. Effectiveness	10	5	50	4	40	1	10
4. Instructional Approach	4	5	20	3	12	1	4
B. Internal Quality							
1. Instruction							
a. Measurable objectives	4	4	16	3	12	2	8
b. Instruction on objectives	10	5	50	3	30	1	10
c. Practice and feedback	6	5	30	3	18	1	6
d. Technical Accuracy	4	4	16	5	20	5	20
e. Interest	4	3	12	4	16	3	12
f. Ease of Use	3	2	6	4	12	2	6
2. Assessment							
a. Frequent checks	6	3	18	4	24	5	30
b. End of Program	2	3	6	5	10	5	10
3. Record Keeping	1	5	5	1	1	5	5
C. Management							
1. Training	1	5	5	2	2	1	1
2. Clear procedures	3	3	9	4	12	4	12
3. Other program interference	2	4	8	2	4	3	6
4. Effect on other programs	2	2	4	3	6	3	6
5. Personnel availability	3	5	15	5	15	5	15
6. Space	6	2	12	4	24	3	18
7. Facilities, equipment, materials	6	5	30	4	24	5	30
8. Logistics, communication time	6	4	24	3	18	5	30
TOTAL	100		411		351		317
Final Average Rating			4.11		3.51		3.17

sophisticated matrices may be constructed, such as those that consider the probably accuracy of the information on which the ratings are based and the confidence of the rater in making judgments.

The advantages, disadvantages and individuals involved in each of these techniques discussed in this section are outlined in the following table.

Technique Summary

Expert's Recommendation	Delphi
Advantages: easy, quick, informal. Disadvantages: not systematic or objective, the least systematic method available. Individuals involved: consultants	Advantages: highly structured, gathers information from experts anywhere, eliminates many biases, can be repeated for ongoing feedback. Disadvantages: time-consuming for total procedure, possible difficulties with mail correspondence, no guaranteed consensus. Individuals involved: consultants

Program Profile Checklists	Systems Approaches (Decision Matrices)
Advantages: objective, structured, requires planners to: specify needs along several categories of information; develop a set of considerations for comparing program. Disadvantages: planners may omit or not agree on some critical program factors; checklist is usually yes/no therefore 2 or more programs may comply, not specific. Individuals involved: (program planners) rating committees,consultants	Advantages: very objective and specific, requires planners to: clarify objectives, alternatives, and related costs; construct model and apply a decision rule to determine the best overall program. Disadvantages: loss of information that can not be categorized, planners may not agree on weights of factors, some individuals may not conceptualize in this fashion Individuals involved: (program planners) rating committee

The next exercise, "Choosing a Program" will give you valuable practice in a particular approach: decision matrices.

SUMMARY

This section outlined some of the ways the evaluator can help the planners formulate procedures for choosing among alternatives. As noted in the introduction, this task will require the evaluator to participate in the planning activities so that appropriate tools can be selected and the evaluator can assist the planners in using them effectively. Some of the tools that might be suggested are hiring consultants, using rating committees, conducting a Delphi study, developing Program Profile Checklists, and incorporating systems approaches, such as decision matrices. It should be remembered, however, that various kinds of program analysis tools may be modified, combined, or even developed to meet the particular needs of a project, and that the evaluator can help planners in determining the most appropriate tool(s) for their needs.

The following exercise, "Choosing a Program," will give you valuable practice in a particular systems approach: decision matrices.

SECTION E

Exercise 4

CHOOSING A PROGRAM

Directions: To aid the program planners in selecting a program from the three programs they had selected as possibilities, Peter Slate helped them construct a decision matrix. The matrix would enable them to compare the relative merits of three different business education programs. He provided the planners with fairly detailed information about the three programs so they could readily compare them and fill in the matrix.

In the following exercise, you will fill in parts of a decision matrix based on brief descriptions of the three alternative business education programs.

Read the program descriptions, keeping in mind the results of the needs assessment conducted in Area E. To review, school officials, parents, teachers and students want a program which meets these criteria:

1. The program is experience-based with students participating in local business life.
2. The program is designed to teach young people to manage money and resources to operate a business.
3. The program is economical.

The program planners have completed most of the weighted scores for Programs A, B and C. Six considerations for each program are still left to do. To fill in the six underlined program considerations in the decision matrix:

1. Assign a weight of 1-5 for each program consideration that has been left blank. Assign a rating of 5 if you feel that a program fully meets a consideration.

2. Multiply each rating by the weight in the left hand column of the table to obtain a score for each consideration.
3. For each program, add the scores of the six underlined program considerations together.
4. Add this total to the subtotals for each program. This will result in a total score for each program.
5. To get a final average rating for each program, divide each total by 100.
6. Recommend one of the programs for adoption in Area E.

Business Education Program A

Business Education Program A proposes to prepare high school students by helping them understand the business world. The program has five major units, each preceded by a statement of instructional objectives, e.g., "to demonstrate a knowledge of various aspects of marketing" and "to understand how to manage the financial aspects of a business."

A Teacher's Manual describes the relationship between the objectives and the instructional materials in each unit. It also includes testimony of business experts as to the technical accuracy of each of the program's units which focus on major business concepts.

A text and workbook designed to introduce basic business concepts comprise the instructional materials for this program. The program also furnishes unit tests and an end-of-program achievement test to measure the attainment of the objectives. Thus, the program does not necessitate the purchase of special equipment and materials. Additionally, the district has sufficient personnel to meet the practical demands of the program.

Concerning the financial features of Business Education Program A, the project requires an expenditure of only $2.00 per pupil for transportation. This minimal cost is possible because the program consists of classroom lectures and brief field trips during which students observe business people in action. Although the program provides no particular training or orientation procedures, it is recommended that teachers be given a one-hour orientation session to explain various aspects of the program: its rationale, how to use the Teacher's Manual, the specific purposes of each unit, and how to administer and score the achievement tests that accompany the program.

The program has been used in several districts. Questionnaire results indicate that 65% of the students considered the program "interesting." Interviews with teachers showed that many employed the program because it fit in with social studies classes. Furthermore, 60% of the teachers felt that the program would probably "be better if students had a chance to participate in business practices in the community." These teachers also found the materials and record keeping procedures "easy to understand and use." More than 70% of the students received a score of 80% or better on the end-of-program test.

Business Education Program B

Business Education Program B has been used, with considerable success, with 5,000 high school students in a large urban school district. Although referred to by one morning paper as an "expensive way to teach thrift," the cost of this program, per pupil, is well within the range of comparable programs. Designed by a team of consumer and business experts, the major purpose of Program B is to enable students to understand the relationship between consumers and the business community. It achieves this end by combining classroom activities with internships in local businesses. Business men and women have indicated their interest in the program by offering to participate in it. Their assistance will help students apply what they learn to real life."

The Teacher's Manual provides guidelines for presenting the concepts, as well as feedback for all activities. The Teacher's Manual suggests that teachers periodically observe the students in the field and interview both the students and the business men participating in the program. Notes from observations and these test scores can be recorded according to the teacher's preferred system. No formal orientation procedure is included in the program. It is not clear how teachers will coordinate these activities. The program, however, does include unit pretests and posttests and an end of the program test.

The instructional materials for students include paperbound books with detachable exercises. Program B requires no additional materials or district equipment for implementation. (The final report describes the relationship between the program's objectives, the instructional materials and the degree of student success. Testimony of subject matter experts about the high quality of the information provided in the materials is also available.) These materials will provide additional practice in using the concepts learned in class and in the field. Students are given numerous classroom opportunities to apply these concepts in a variety of simulations and written exercises.

Student achievement and attitudes were good - more than half of the students received a score of 83% or more on the end-of-program test that comes with the program.* In addition, questionnaire results indicate that 8% of the students thought that the program was "interesting and enjoyable."

---

*The coprehensive combination of "real-life" practice, class room instruction and written material seems to account for this high percentage.

Business Education Program C

Business Education Program C is designed to enable students to understand the relationship between consumers and the business community. "Understanding" occurs through students' active participation in the community's businesses and banks. The program employs broad, nonmeasurable goals such as "to learn practical consumer strategies" and "to deal effectively with consumer problems."

Program C emphasizes practical experience. During the tryout of the program, a cross-section of business men and women participated in the program. According to the final report, over 25 business people originally agreed to take part in the program. Eight dropped out claiming (in follow-up interviews) that they had lost touch with what was happening in the program. However, local business men and women still express an interest in participating in the actual program.

At a cost of $10 per pupil for training materials, Business Education Program C includes an end-of-program measure and an interview schedule to use with members of the business community who participate in the program. To further enhance the program's feasibility, it does not require the use of the district's already scarce audio-visual equipment nor does it require classroom time. Time requirements for various activities are clearly stated. The program also provides for the training of business persons to explain the rationale behind the program, how to use the Business Person's Manual and how to implement the program. Eighty percent of the business people indicated that the procedures for their training and for keeping records were "easy to use."

The program includes an individualized text designed to reinforce concepts students learn during their practical experience. Eighty percent of the business ed teachers rated the caliber of the materials as excellent. The text presents students with a variety of problem situations involving businesses and students and suggests appropriate actions that business people might take to solve the problem. However, care must be exercised in determining the connection between the program's activities and its objectives.

The program was tried out in an urban high school district with over 5,000 students. An end-of-program test showed that 90% of the students received a score of 55% or more on a 50-item test. On specially prepared questionnaires, 80% of the students reported that the program was "interesting and enjoyable."

FINAL RATINGS OF ALTERNATIVE BUSINESS EDUCATION COURSES

PROGRAM CONSIDERATION	Weight	Program A Rating*	Program A Score	Program B Rating*	Program B Score	Program C Rating*	Program C Score
A. GENERAL CHARACTERISTICS							
1. Goal Agreement	10	2	20	5	50	5	50
2. Cost	7	—	—	—	—	—	-
3. Effectiveness	10	4	40	4	40	2	20
4. Instructional Approach	4	0	-	4	16	5	20
B. INTERNAL QUALITY							
1. INSTRUCTION							
a. Measurable objectives	4	—	—	—	—	—	—
b. Instruction on objectives	10	3	30	5	50	2	20
c. Practice and feedback	6	2	12	5	30	2	12
d. Technical accuracy	4	4	16	5	20	3	12
e. Interesting	4	3	12	4	16	5	20
f. Ease of use	3	5	15	0	-	5	15
2. ASSESSMENT							
a. Frequent checks	6	—	—	—	—	—	—
b. End-of-program	2	—	—	—	—	—	—
3. RECORD KEEPING	1	5	5	0	-	5	5
C. MANAGEMENT							
1. Training and orientation	1	—	—	—	—	—	—
2. Clear procedures	3	0	-	3	9	3	9
3. Other program interference	2	0	-	0	-	0	-
4. Positive student performance on other programs	2	2	4	0	-	0	-
5. Personnel availability	3	—	—	—	—	—	—
6. Space	6	0	-	0	-	0	-
7. Facilities, equipment and materials	6	4	24	3	18	5	30
8. Logistics, communication, and time	6	0	-	0	-	4	24
Subtotal			178		249		237
Total of 6 considerations to be filled in							
Final Total							
Final Average Rating							

*Ratings: 5=Very Desirable; 4=Desirable; 3=Average; 2=Undesirable; 1= Very Undesirable; 0=Not present

BUILDING IN THE EVALUATION PLAN

SECTION F

EVALUATION QUESTIONS

Section F

EVALUATION QUESTIONS

F. Evaluation Questions

1. Question Categories
2. Formulating Questions
3. Exercise 5

The planning staff and Peter Slate, the evaluator, have finally chosen a Business Education Program. Now they are ready to discuss the evaluation of the program. Since Peter Slate has been working with the planning staff since its inception, he is able to help the staff with decisions on the purpose and scope of the program's evaluation, before the program is even implemented. These early evaluation plans, though desirable, might not always be possible. An evaluator may not be called in until after the program is implemented, or; unfortunately, sometimes not until after the program is completed.

Question Categories

Regardless of when the evaluation plans are made, the evaluation should include the following categories of evaluation questions in order to conduct an effective evaluation and ensure that the data to be collected yields the desired information. These categories are only meant to help the evaluator and program planners organize their evaluation questions. They are arbitrary and there is some overlap between categories.

1. Outcome Questions: Did the program meet its objectives? (cognitive, affective)
2. Implementation Questions: What does the program look like in operation? Is it being implemented as planned?
3. Cost Questions: What costs, including $ expenditure, time, and resources are necessary for the program?

4. Research Questions: How do specific program processes relate to program outcomes?

5. Unanticipated outcome Questions: What other outcomes, positive or negative, can be attributed to the program?

Evaluation questions are important because they give direction to all of the activities that follow. They guide the selection of data collection instruments, evaluation design, and data analysis procedures. They help increase the efficiency of the evaluation by identifying the kinds of information that must be collected and the kinds of information that are superfluous to the needs of the decision makers. Each category of question is discussed in more detail below.

1. Outcome Questions: Did the program meet its objectives?

For many programs, most or all of the program objectives are concerned with people and the way the program affects their behavior, knowledge, feelings, and so on. (Some programs do have objectives that are not related to changes in people e.g., cost factors; how an evaluator should treat such "non-people" objectives is discussed in a later section of the Workshop.)

— Depending on the objectives of the program, evaluation questions can focus on any group of individuals (e.g., students, parents, teachers or aides) who are expected to be affected by the program.

Typically the objectives of a program will describe such outcomes as

— the major types of skills or competencies that people are to achieve

— the areas of content they will learn and/or

— the attitudes, values or feelings they may acquire as a result of the program

2. Implementation Questions: Is the program being implemented as planned? What are the key characteristics of the program? What does the program look like in operation? What are program processes or procedures?

Documenting the program's implementation is necessary for program certification and program replication. If program goals are achieved, and it seems that the outcomes were caused by the program, implementation information will suggest which procedures caused which outcomes. A major goal, in fact, of many agencies is to identify successful programs and replicate these programs in other settings -- by definition replication calls for procedures to be known so they can be duplicated.

To determine if a program is being implemented as planned, evaluators collect two kinds of information: background information about the program and information about program processes.

A. Background Information: What are the key characteristics of the program? The evaluator collects most of this information while investigating the context in which the program operates. (Discussed in Section D)

1. Origin of the program
2. Program Objectives
3. Individuals served by the program
4. Individuals operating the program

B. Program Processes: What are the program's processes? What does the program look like in operation? Evaluators need to be selective when they ask questions about processes. Although processes are the means by which goals are supposed to be met, it would be senseless for an evaluator to investigate the hundreds of things that people do while a program is going on.

Therefore, it is suggested that evaluators focus on those processes that are expected to have the greatest influence on planned outcomes, are unique to the program, or are most costly (in terms of time and money).

The term "process" covers a wide range of things. For instance, it refers to:

- the activities of students, teachers, and others involved in the program
- instructional strategies employed in the program
- materials and resources used by program participants
- basic program procedures.

By comparing information about program processes during its use with those processes outlined in the program plan, evaluations are able to determine the extent to which the program is being put into operation as planned. This information is important to document its procedures and modify them if necessary.

3. Cost Questions: What costs including dollar expenditure, time, resources are necessary to plan and implement the program?

"Cost" is a word that takes on different shades of meaning depending on the persons using it. For example, cost can refer to:

- dollar amounts spent on materials, salaries, construction and other budget items
- allocation of personnel
- wear and tear on program staff.

Cost-type questions about program elements should not be asked in the absence of research questions that relate program components to outcomes. Decision-makers are likely to treat a costly program component differently

if they know that it is crucial to the achievement of program objectives. Cost questions are, by definition, tied to decisions about the value of a component, or even an entire program. These decisions can become less subjective when evaluators systematically research the relationship between specific program characteristics and outcomes and report the results.

4. Research Questions: How do specific program processes relate to program outcomes?

Research questions involve a planned variation in the program processes to determine what differential impact, if any, the variation has on attainment of program outcomes. For example, a decision-maker may want to know whether small group discussions or individual workbooks are more efficient for producing certain program objectives. Evaluators should ask research questions about those program processes that they expect will have the greatest impact on outcomes, are unique or costly in terms of time, money or resources. The critical characteristic of research questions is that all other program conditions remain the same while some single program component is systematically varied.

5. Unanticipated Outcome Questions: What other unplanned outcomes can be attributed to the program?

This question calls for an identification of outcomes -- either positive or negative -- that can be attributed to the program but which were not a part of the original program objectives. Considering all program outcomes gives decision makers a more thorough picture of just what the program does accomplish than would an evaluation that measures only planned outcomes.

During the operation of any education program, there is always the possibility that outcomes may occur which were not anticipated, but that nevertheless have great impact. For example, students may master cognitive objectives, but may develop an aversion toward the subject matter. Unanticipated outcomes may be positive or negative. They may be limited to the program being evaluated; they may affect other programs; or, they may affect the school program in general.

Sleuthing for unanticipated outcomes or additional information may pay off in significant findings, but there is also the risk of time being spent on the collection of data that is either unrelated to the program or related but fairly trivial. Just how much digging around is done in search of the unexpected will be determined by the purpose and scope of the evaluation. For example, a time-constrained evaluation for the purpose of making a year-end report to the school district allows much less flexibility than a longer evaluation in which one of the concerns of those commissioning the evaluation is to discover all of the outcomes and spin-offs that can be reasonably attributed to the program in question.

The process of determining whether the planned program objectives have been achieved is a sophisticated process involving selection or development of valid and reliable measures and a research design that controls for as many extraneous variables as possible. Any leads which seemingly point to unanticipated program effects must be as thoroughly and systematically pursued as the planned outcomes to avoid the pitfall of making claims about the program that are based on partial evidence.

Formulating Evaluation Questions

The evaluation questions which are particular to each program should be arrived at jointly by those commissioning the evaluation (or their designates) and the evaluator. This assures that the questions will be consistent with the information needs of the decision-makers and compatible with whatever resource and social constraints affect the evaluation.

The language used to phrase evaluation questions must be understood by the decision-makers and others involved in the evaluation. Ideally, everyone involved will be able to look at a list of evaluation questions and agree on what the questions mean. Clearly stated questions will help the evaluator determine what types of information should be collected, when, and from what respondent group(s).

An evaluation question can be stated with a criterion (e.g., "Did the program cost per student stay within $10.00 of the cost of the regular program?") or can simply be a call for general information (e.g., "How much did the program cost per student?"). Whether or not criteria are included in the evaluation questions will largely be a matter of whether or not the goals of the program are stated in operational terms.

The precise number of evaluation questions is limited by the amount of relevant information decision-makers need and the constraints within which the evaluation must operate (e.g., time, money, resources, and personnel). There may have to be some compromises between the evaluator and groups concerned with the program and the evaluation in order to keep the length of the list manageable.

Below is a partial list of the evaluation questions Peter Slate and the program planners have formulated: they are the result of a session in which program planners carefully examined the experience based Business Education Program to determine appropriate evaluation questions.

1. Outcome Question

   Are students able to list the responsibilities of both a business person and a consumer in a business transaction?

2. Implementation Questions

   Are businesses cooperating in the program?

   Is a one hour orientation session being arranged? Is it being attended by parents, businessmen, teachers and students?

3. Cost Question

   What is the effect of the internships on other class work?

4. Research Question

   Does the type of business used in the internship affect achievement of course objectives?

5. Unanticipated Outcome Question

   Does the course experience affect career choices?

Since this is the first semester that the new business education program will be implemented, the program planners want the evaluator to collect a great deal of data concerning the way in which the program is being put into operation. This kind of evaluation which is being held at Kennedy High is a formative evaluation in which implementation questions are of prime concern.

These same categories of evaluation questions however, would be used in a summative evaluation though the emphasis, would probably not be on program implementation. Ideally, a summative evaluation is conducted only when a program has passed its developmental stage and the implementation factors need only be used as background information.

Since Peter Slate is involved in a formative evaluation he will help plan and check on the implementation of the program, whether it is running according to plan and on schedule, whether the right students are getting the instruction as prescribed and so on. Frequent checks of student progress and other program objectives also are made. All of this information is periodically reported by the formative evaluator to the program staff -- so they can correct any problems and know exactly what is happening. In short, formative evaluations provide the decision-makers, in this case the people running the program, with feedback during the evaluation effort in order to help them improve the program.

The decision makers in Area E are concerned with a formative evaluation of the Business Education program this coming year, but they are also planning a summative evaluation of the program for the following year. Admirably, this is the best way to conduct evaluation, a formative evaluation for the first year in order to give feedback and assist in the proper implementation of a program and then a summative evaluation in another year once the program is functioning as designed. The summative evaluation will then compare the new Business Education program with the old Business Education program. This summative evaluation information will be used to decide whether or not to continue the program, expand it, or whether to discontinue it.

Section F

Exercise 5: Formulating Evaluation Questions

In this exercise you will use the partial list of evaluation questions Peter Slate and the planning committee wrote as guides to formulate other evaluation questions which might be asked in the evaluation of Experience-Based Business Education.

Refer to the feedback from Exercise #3 Determining the Program Planning and Evaluation Context. This information will help you formulate evaluation questions in this exercise.

A. Program Outcomes

1. Are students able to list the responsibilities of both a business person and a consumer in a business interaction?
2. _____

_____

_____

B. Implementation Questions

1. Are local businesses cooperating in the program?
2. _____

_____

_____

C. Cost Questions

1. What is the effect of the internships on other class work?
2. _____

_____

_____

D. Research Questions

1. Does the type of business used in the internship affect achievement of course objectives?
2. _____

_____

_____

E. Unanticipated Outcome Questions

1. Does the course experience affect career choices?
2. _____

_____

_____

Section F

Feedback

Exercise 5

Formulating Evaluation Questions

Outcome Questions

1. Are students able to list the responsibilities of both a business person and a consumer in a business interaction?
2. Are students able to manage money and resources to operate a business?

Implementation Questions

1. Are the local businesses cooperating in the program?
2. Are the students receiving instruction outside the classroom as interns in the local businesses?
3. Is a one hour orientation meeting being arranged? Is it being attended by parents, businessmen, teachers and students?
4. Are the internships being scheduled so as to not interfere with attendance in other classes?

Cost Questions

1. What is the effect of the internships on other class work?
2. What can be done to keep down the cost of the internships?

Research Questions

1. Does the type of business used for internship affect achievement of course objectives?

Unanticipated Outcome Questions

1. Does the course experience affect career choices?
2. Will the program affect relationship between the business community and the students?

BUILDING IN THE EVALUATION PLAN

SECTION G

MEASUREMENT CONSIDERATIONS

SECTION G

MEASUREMENT CONSIDERATIONS

Now that a set of evaluation questions has been identified (i.e., outcome, implementation, cost, research and unanticipated outcome questions), Peter Slate's next task is to select the means by which these questions will be answered. The role of measurement is indeed critical, since the measurement instruments underlie the technical credibility of both formative and summative evaluation results. The relationship between poor instruments and the results they yield can be summarized by a phrase borrowed from computer programming, "garbage in--garbage out!"

There is a rich and sophisticated technology associated with the construction of good test and measures, but an evaluator need not be an expert psychometrician to select credible methods to gather data and avoid the "garbage syndrome." There are several guidelines, described below, that evaluators should follow with all data collection instruments:

1. Make sure the measurement device is directly linked to the evaluation question for which it is being used.
2. Give consideration to alternative measurement strategies.
3. Make sure each measure is appropriate for its recipients.
4. Measurement instruments should be as valid and reliable as possible.
5. Make the data collection as unobtrusive as possible.
6. Instruments should be sensitive to positive as well as negative outcomes.

1. Make sure the measurement device is directly linked to the evaluation question for which it is being used.

There is always the temptation to collect information merely because it might prove "interesting," or to use a particular measurement device because it is routinely administered at a school site. Unless these also promise to be of significant help in answering evaluation questions, they should not be used for evaluation purposes.

In short, evaluators should carefully examine all potential evaluation measures -- especially published instruments -- to be certain that there is a match between the measure and the specific information needs. For example, a measure labelled "Grade Two Basic Math Computation Test" might contain many word problems, whereas the evaluation question called only for assessment of simple numerical calculation skills (10x10=_____). Mismatches like this one between what an instrument measures and the intent of evaluation questions, can be unfair to the program and can also damage the credibility of the evaluation. Even if an instrument is entirely suitable in content, it is nonetheless possible for the results to be ambiguous or misledading. This problem is becoming increasingly apparent in achievement testing and has been captured in the debate over the relative merits of criterion-referenced tests and norm-referenced tests.

The distinugishing feature of good criterion-referenced tests (CRTs) is that they are designed to measure clearly delineated skills (intellectual and motor), knowledge, or abilities (mental or physical). Results from CRTs indicate an examinee's level of achievement for each area assessed in the test. In other words, the results show an individual's level of mastery for specific skills or abilities.

In contrast, norm-referenced tests (NRTs) compare an examinee's performance with that of a large group of other people, the "norm" group. Results from such measures indicate a person's relative standing in comparison to the norm group. Scores often are expressed as percentiles, grade equivalent scores or standard scores. NRT results provide relative scores describing how a person performed in a general skill area, such as reading comprehension.

When mastery of specific skills or knowledge is to be inferred from test performance, criterion-referenced tests are preferable. They can be obtained from measurement specialists who develop CRTs (also called "domain-referenced tests" or "objectives-based tests"), or they can be developed locally by school personnel.

If data comparing individual students to an "average" student's scores are needed, then a test which provides relative scores should be used. If a norm-referenced test is to be administered, be certain that the group on which the test was normed is comparable to the group which is to be tested; if not, the results may be unfair. Information on the norm group is usually published in the test manual.

It is important to note that these comments are not limited to achievement measurement devices. They apply to <u>all</u> such devices.

2. <u>Give consideration to alternative measurement strategies</u>.

Because evaluations generally are concerned with assessing a full range of program characteristics including activities and materials, time and costs, attitudes, and, of course, achievement, a range of measurement approaches may be needed. Pencil-and-paper measurement instruments, such as the traditional achievement-type test or quiz, are by no means the sole method for collecting evaluation data. Options open to

evaluators, depending upon the nature of the evaluation question, include such strategies as observations, interviews, questionnaires, performance tests, records, checklists and so on; such strategies can be used to obtain information about and from students, teachers, parents and other sources. Evaluation findings are strongest when based on corroborating evidence from several different instruments, especially when different types of instruments (observation and interview) are used with different respondents (students and teachers).

3. Make sure each measure is appropriate for its recipients.

Whatever measurement devices are selected, each device should be appropriate for those taking it. For example, an interview would obviously be more appropriate for non-readers than a questionnaire. Again, the credibility of results can be challenged if it appears that the method of measurement was inappropriate for the situation (e.g., speeded test is inappropriate when program goals did not put objectives into any time frame).

4. Measurement instruments should be as valid and reliable as possible.

The terms "validity" and "reliability" are used frequently by test and measurement specialists, and for a good reason: validity and reliability are critical attributes for all measurement instruments, whether the instrument is a clock, a high powered electronic sensor, a personality inventory, or a reading test. The concepts of reliability and validity can take on relatively complex technical and statistical meanings when they are applied to the process of test construction, a topic beyond the scope of this workshop. However, they also have fairly common sense interpretations with which evaluators should be familiar when choosing measures.

Reliability refers to the consistency with which a measure assesses whatever it is intended to measure. One way a measure can be reliable is over time. For example, a group of students might take an algebra test on Monday. Provided that they receive no practice or instruction on the class of problems included in the test, the group should receive approximately the same score if the same test were to be administered again a week later. Such a test would have demonstrated reliability, or stability, over time.

Another type of reliability, internal consistency, is considered an estimate of the degree to which a group of test items are performing their functions consistently. A test that shows a high level of internal consistency indicates that items on the test, or in a sub-test, are working together to measure the same thing, rather than many different things.

These considerations lead naturally to a question of validity, that is, what "thing" or "things" does a particular instrument actually measure? While reliability reflects the consistency of a measuring device, it is *validity that indicates whether a test measures what it is supposed to measure.* Though the presence of reliability makes it <u>possible</u> for a test to be valid, it does not guarantee the test's validity. For example, it is possible to have a bathroom scale that is consistently five pounds overweight; the scale is reliable--consistent--but is not a valid indicator of true weight.

In summary, when an instrument has demonstrated a high degree of reliability, an evaluator can be reasonably sure that it will measure consistently with different groups and at different times, which is especially important when an evaluation design calls for pre and post-testing or the use of comparison groups. When a test has proven validity,

it can be trusted to measure accurately what it is intended to measure. The concept of validity depends upon the use to which a measure will be put. A test could, for instance, be a valid measure of mechanical reasoning, but if an evaluation is to assess mechanical ability, then the test is not appropriate. Note that a test that is not appropriate may still be valid a measure for the skill area it does measure.

As a final note, data on the reliability and validity of most published tests are usually available in the tests' manuals. Unless evaluators are fairly confident in their ability to interpret test manuals, which can be baffling indeed, it is advisable to seek the aid of a measurement specialist or test review books. Determining the extent to which tests developed by local school personnel are valid and reliable is slightly more complicated. Few local sites have the resources to conduct the required time consuming and expensive validity and reliability studies. Nonetheless, evaluators should be prepared to justify the use of locally-developed instruments, perhaps asking program personnel to review the instruments. Again, when in doubt, the assistance of a person trained in educational measurement would be of value.

5. Make the data collection as unobtrusive as possible.

All measurement alternatives should be implemented in a way that does not excessively intrude into the program's normal operation. Ideally, measurement devices are useful for instructional as well as evaluation purposes. The ultimate purpose of evaluation is to improve the quality of education for students, parents, teachers, and others, but this purpose can be defeated when the evaluation effort inordinately interferes with the learning/teaching process.

6. Instruments should be sensitive to positive as well as negative outcomes.

In addition to picking up the positive outcomes that a program is designed to effect, measures should also be constructed to pick up information on possible detrimental program effects.

BUILDING IN THE EVALUATION PLAN

SECTION H

CHOOSING TESTS TO MEASURE ACHIEVEMENT OUTCOMES

CHOOSING TESTS TO MEASURE ACHIEVEMENT OUTCOMES

H. Choosing Tests

1. Types of Instruments
2. Test Specifications

Types of Instruments

Peter Slate's next task is to choose measurement instruments for each of the evaluation questions he and the program planners have identified. Table 1 on the following page outlines the most commonly used measures: tests, reports, observations and records.

Since this is the first time the new program is being put into operation, Slate will conduct a formative evaluation. He is concerned with frequently monitoring the program's implementation and measuring students' progress in attaining the program's enroute and terminal objectives. By contrast, if he were conducting a summative evaluation he would be concerned with documenting only major activities in the implementation of the program. Similarily, students' attainment of the program's objectives might be measured upon completion of major program units and at the end of the entire program.

Generally a program's objectives can be placed in two categories:

1. Cognitive and Psychomotor Objectives (achievement or skill outcomes)
2. Affective Objectives (attitudes, feelings)

Cognitive and Psychomotor objectives are often critical areas of concern in an evaluation and this section of the workshop will deal with these objectives. The next section of the workshop, Section I, will cover measurement of both the Affective Objectives and the remaining evaluation questions which are important in a formative evaluation.

Although the Experience Based Business Education Program that has been selected, Program B, includes both interim and end of the program achievement tests, evaluators will frequently have to select or develop achievement tests to measure each learning outcome in a program.

Table 1 - Measurement Devices

I. Tests

A. Selected response tests (e.g., true-false, matching, multiple choice)

B. Constructed response tests

1. "question-and-answer" tests (e.g., essays, short answer, completion items)

2. work samples (e.g., examinee produces a product - paints a picture, sews a shirt, completes a job application form, writes a business letter, etc.)

C. Process tests (e.g., a performance is judged as it occurs - reading aloud, playing piano, throwing a softball)

II. Reports

A. Questionnaires and surveys

1. "choice" items (e.g., rating scales, ranking scales, semantic differential)

2. open-ended questions

B. Interviews

C. Logs and journals

D. Formal reports (examples: annual reports submitted by a principal to the district office, status reports prepared by a program director)

III. Observations

A. Structured (example: a checklist of specific behaviors which an observer uses to record student actions)

B. Unstructured (example: a form on which a teacher periodically records impressions of the way classroom activities occur)

C. Inventory (example: a checklist used by the school librarian to inspect program materials)

IV. Records (including all regularly maintained school records such as attendance records, minutes from meetings, cafeteria sales; classroom records indicating use of equipment, student progress, etc.)

Test Specifications

No official tally has been made of the number of program staff members who have exclaimed upon reviewing disappointing test results -- "But the test didn't assess what we taught!" Nor has any record been kept of the decision-makers who have asked, "What does this really mean?" when presented with evaluation findings such as, "70% of the students achieved a score of 75% on the XYZ Test." Such concerns, though, are real, and they have highlighted the need for tests that are sensitive to instruction and which provide meaningful information about achievement of cognitive and psychomotor outcomes.

One approach to selecting or developing appropriate achievement tests is to outline, for each learning outcome to be measured, a set of test specifications. A test specification is a clear definition of the examinee behaviors to be measured and the context in which they will be measured.

There are two basic ingredients in a set of test specifications:

1) a description of the content, that is, what do examinees encounter before making a response?
2) a description of the kind of response required of examinees, i.e., what are they expected to do during the test?

Making test specifications may at first seem to be the long way around for accomplishing the apparently straightforward task of choosing valid and reliable tests, but it is a useful procedure for several reasons:

-- The process of detailing the stimulus and response elements can help ensure that both program staff and the evaluator share the same interpretation of the outcome to be measured. The fairness of the evaluation clearly is maximized when there is a match between what was taught and what was tested.

-- On the practical side, test specifications give the evaluator guidelines for reviewing the suitability of existing instruments

-- When a test is to be developed by the evaluator or staff, the specifications can serve as a "blueprint" for generating test items.

-- Clearly written test specifications help to remove the mystery from testing. They provide a way to make the evaluation open to public inspection by allowing staff, parents and others to see exactly what was measured.

-- The specifications can be used as a vehicle for communicating with decision-makers who will interpret examinee achievement on the basis of test scores.

The contents and format for test specifications can vary from complex to simple. As a rule of thumb, more detail will be needed when the specifications are to be used as a basis for developing tests rather than just reviewing tests, since the specifications must be complete guidelines for the test item writers. Above all, test specifications will effectively communicate to staff, test writers, decision-makers and other groups who have a stake in the evaluation exactly what is to be measured and in what context. Some of the questions that might be considered when developing test specifications are outlined on the next page, followed by a test specification "sample."

CONSIDERATIONS IN MAKING TEST SPECIFICATIONS

I. STIMULUS ELEMENT OF THE TEST: WHAT DO EXAMINEES ENCOUNTER BEFORE MAKING A RESPONSE?

A. Test content

- What is the range of subject-matter to be covered in the test?
- What concepts, rules, principles should be incorporated into the test?
- What skills or abilities are to be measured?
- What are the limits of the content of the test?

B. Test format

- How should the test content be presented?
- What should the test items look like?
- What directions should be given the examinees?

C. Materials

- What materials, equipment, or supplies are needed for the testing procedure?

II. RESPONSE ELEMENT OF THE TEST: WHAT ARE EXAMINEES EXPECTED TO DO?

A. Form of the response

- Should examinees choose an answer from among options given on the test ("selected response" items): multiple choice, true-false, matching, etc.?
- Should examinees construct their own responses?
  - Should they respond orally or in writing to questions? problems? statements? (examples: essay tests, short answer, completion items)

- Should they provide a "sample" of work? (examples: sew a blouse, fill out a job application form, make a wood construction)
- Should examinee's performance be judged as it occurs? (e.g., dancing, reading aloud, throwing a softball)

B. Criteria of Correctness

- What constitutes a correct response?
- What constitutes an incorrect response?
- For selected response items, what are the "wrong answer" options?

III. ADMINISTRATION OF TEST

- How much time is permitted?
- Do administrators need special training?

Sample Test Specification

EVALUATION QUESTION: Can the student discriminate between assumptions and reliable observations regarding a given picture?

I. Stimulus Element

A. Content

Learner will be given a simple drawing depicting several individuals in potential interaction. The drawing may depict any scene whose symbols would be familiar to the learner population. The drawing should contain elements that are factually observable and aspects that are open to interpretation. The latter aspects should suggest a conclusion.

Following some of the drawings, there will be 5-10 simple statements about the objects and/or actions in the picture. Statements may refer to any individual or object in the picture, or to any group within the picture.

Statements will be of two types:

1. assumptions, or
2. observation statements

B. Test Format:

Directions: "Study each drawing, and the statements which follow. Decide whether each statement is a reliable observa-

tion statement about the drawing or whether it is based on an assumption. Write the letter of your answer on the answer sheet."

1. The girl is coming to assist the boy.
   a. reliable observation statement
   b. assumption
2. The girl is running towards the boy.
   a. reliable observation
   b. assumption
3. The boy is injured.
   a. reliable observation statement
   b. assumption
4. The boy has been attacked by the dog.
   a. reliable observation statement
   b. assumption

C. Materials

Test book for each student.

II. Response Element

A. Form of response

Learner writes the letter of his choice on a separate answer sheet.

B. Correct response

Assumption statements should be based on conclusions that could easily be drawn from the picture -- i.e., regarding feelings of individuals, potential actions, or guesses about what is transpiring in the picture.

Observation statements should reflect only what is overtly observable in the picture.

III. Administration

15-20 minutes is needed

SECTION I

MEASUREMENT DEVICES

SECTION I

TYPES OF MEASUREMENT DEVICES

I. Types of Devices
   1. Tests
   2. Reports
   3. Observations
   4. Records

An approach to selecting an/or developing achievement tests to answer questions about cognitive and psychomotor outcomes was outlined in Section H. Clearly, achievement tests will not yield all of the information required in an evaluation, and evaluators will need to consider additional measurement strategies to answer other evaluation questions and affective outcomes that are important in both formative and summative evaluations.

Since the first major category of measurement devices, tests, has already been discussed in the previous section, this section will describe the remaining devices listed in Table 1: Self-reports, observations and records.

Some major concerns in the development of reports and observation schedules will be discussed. These are meant to be representative rather than exhaustive.

I. Tests (covered in Section H)

II. Reports

A. Questionnaires

Questionnaires are an economical way to gather information from a wide audience since they can be mass produced and mailed to many people. If they are carefully constructed to provide useful information, they are valuable measurement devices. It is not easy, however, to develop a good and useful questionnaire. The content of a questionnaire is important because it must gather information the decision makers need. Its format is also important because it should motivate people to complete the questionnaire. A good strategy is to develop a clear, short questionnaire. It can include "open-ended" questions which ask the respondent to write in answers and selected response or "closed" questions which present a series of choices from which the individual checks off or circles a choice. The following example illustrates a combination of the two methods, which is perhaps the most fruitful tack to take.

1. "Have you noticed the amount of violence on T.V.?"

Yes 1

No 2

Don't know 3

If "yes"

2. "How do you feel about the amount of violence on T.V.?"

Approve 1

Disapprove 2

Undecided 3

3. "Why do you feel this way?" _____

_____

_____

(Example suggested by A. N. Oppenheim's Questionnaire Design and Attitude Measurement, Chapter 3).

While open-ended questions are easier to develop, the information is harder to analyze because of the wide variety of answers they can elicit. In contrast, closed questions are more difficult to construct but they are easier to analyze.

Before constructing the questionnaire, it is wise to conduct open-ended interviews with some program participants to determine some of the objectives that need to be measured or explored. A rough draft of the questionnaire should then be developed based on these interviews. This draft is given to a small group of program participants to see if the questions are finding out what the evaluator wanted to know and if the questions are clear enough that respondents are able to answer similar questions consistently. For example, a questionnaire may contain two questions asking whether students felt that a segment of instruction was clearly presented. If the answers to these questions are not similar, there may be a problem with the way the questions are phrased.

B. Interviews

Interviews are an alternative to questionnaires and are especially useful when the population to be questioned is fairly small. They are more expensive to administer than questionnaires because of costs associated with hiring and training interviewers. However, when administered

by a trained interviewer they have certain advantages over questionnaires. The interviewer is able to go beyond the list of questions and collect other information that would not show up on a questionnaire. The interviewer can also probe responses by asking followup questions. However, it is important that the interviewer cover the same list of questions with each person interviewed. This consistency will ensure valid results and allow a comprehensive summary of the interviews.

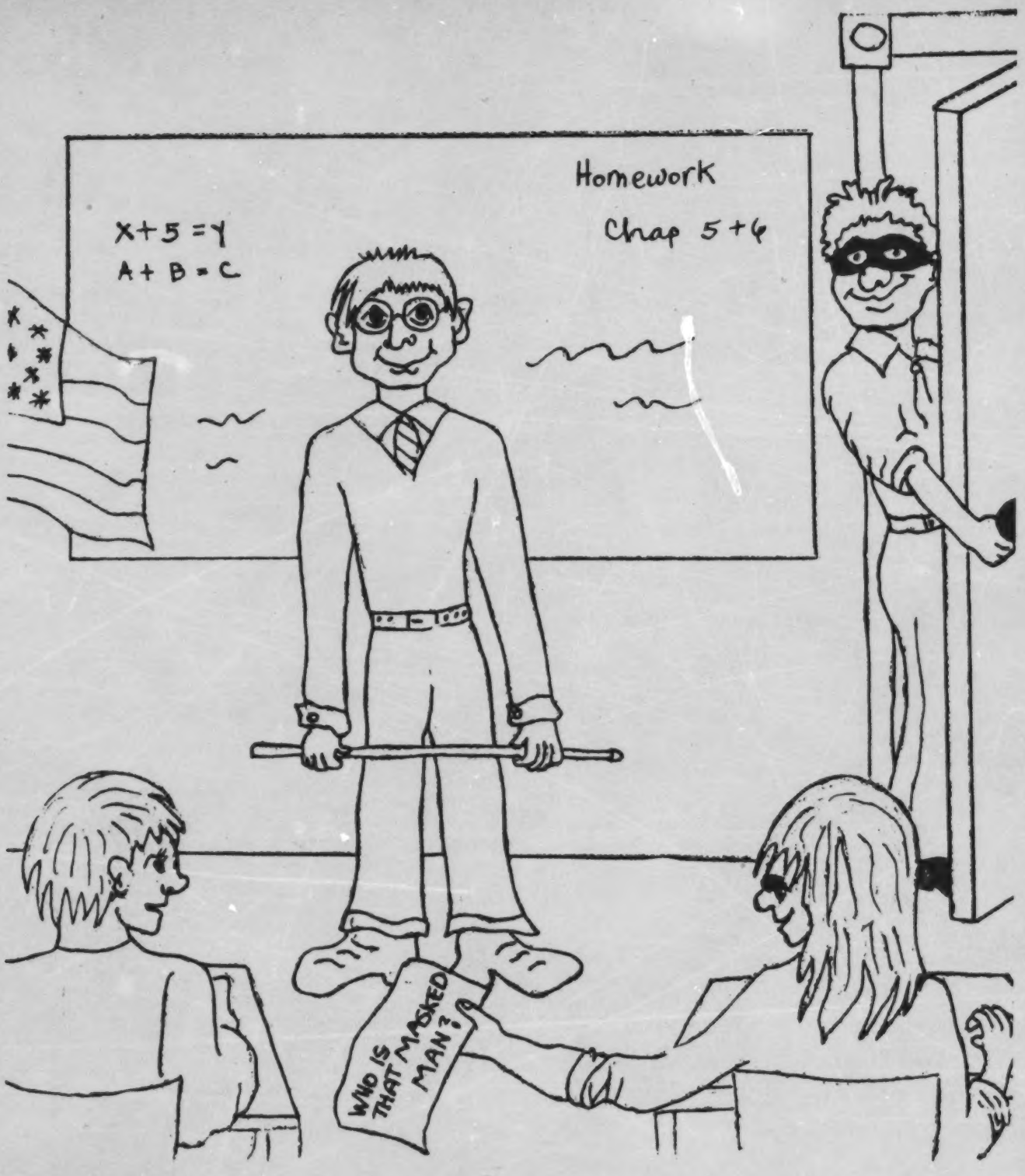
III. Observations

Observations in evaluation can be informal or formal. Informal observations involve a casual look at a program by a few observers. They are quick and easy to carry out and are especially useful in small scale evaluation. Formal observations are more strictly controlled. They specify times of observations and methods for recording information.

Observations can take place in either "controlled" situations where people know they are being observed or in "natural" settings where people are observed in everyday situations and measurement is unobtrusive. In natural settings the evaluator does not manipulate conditions. There is a good possibility that the evaluator's presence is unnoticed, or if noticed, should not affect his findings since he is not having any direct interaction with those being observed. The evaluator can examine samples of behavior and pieces of conversation in such varied settings as assemblies, faculty and student lunchrooms, playgrounds, faculty meetings, PTA meetings, and recesses. For example, attending faculty and PTA meetings could help identify the kinds of praises, doubts, and constructive recommendations that exist about the program that is being evaluated.

IV. Records

Carefully kept school and class records can be one of the most substantive yet inexpensive sources of information. However, even when the records appear accurate and complete, an evaluator should try to utilize other measurement devices to substantiate the findings. For example, an evaluator ran into this situation midway through a Summative Evaluation of a Counseling Program. She discovered that many eleventh graders involved in the program were being reported absent from their study periods. Further investigation from interviews with students and teachers revealed that these same students had independently organized specific subject study groups during that period of the day. This gave a very different picture than the absenteeism records.



Who is that masked man? _____

_____

(Feedback is on page I-7)

(Feedback: The unobtrusive evaluator)

Once an evaluator is familiar with the various measurement devices, it is necessary to decide which device is needed for each evaluation question. Chart 1 below suggests those devices that are usually most appropriate for measuring each type of question.

Chart 1

Measurement Devices Commonly Used for Evaluation Questions

Evaluation Questions	Suggested Measurement Instrument
1. Outcomes	
A. Cognitive and Psychomotor outcomes	Achievement and performance tests
B. Affective Outcomes	Questionnaires, Interviews, Observations
2. Implementation	Records, Observations, Questionnaires, Interviews, Reports
3. Cost	Charts, Questionnaires
4. Research	Tests, Questionnaires
5. Unanticipated Outcomes	Questionnaires, Interviews

MEASUREMENT CONCERNS FOR EVALUATION QUESTIONS

The chart on the previous page gave a brief summary of often used measures for each evaluation question. However, there are particular characteristics for each evaluation question that will influence the type of measure that would be appropriate to use. These characteristics are discussed on the following pages.

Evaluation Questions

1. Outcomes
   A. Cognitive and Psychomotor
   B. Affective
2. Implementation
3. Cost
4. Research
5. Unanticipated

1A. *Cognitive and Psychomotor Outcomes* (discussed in Section H)

1B. *Affective Outcomes*. Measuring affect -- attitudes, opinions, feelings, beliefs -- is not an easy task. Unlike a cognitive or psychomotor outcome (running the fifty yard dash, multiplying two digit numbers), a person's affective orientation is a state of mind. In other words, it is impossible to directly tap a person's feelings, opinions, or beliefs for measurement purposes. Instead, an evaluator can only infer that people hold certain attitudes through their words and actions, and only under particular conditions and situations. With these concerns in mind, what can an evaluator do to measure important affective outcomes?

In all instances, the evaluator, staff, and other interested groups need to determine what they will accept as "proof" of the attitude. Consider an affective outcome question: "Did students' attitude toward school become more positive?" Is it enough for students to say in a

questionnaire or interview that they "like going to school?" Or would it be more convincing to determine if they voluntarily attended an after-school learning lab?

In most instances that involve measurement of attitudes, the evaluator's wisest course of action is to use at least two different measurement strategies. If results from multiple measures aimed at assessing the same affective state of mind are consistent, then the evaluator can have much more confidence in the results. If results are contradictory, then perhaps one -- or all -- of the measures are not adequately tapping the attitude, and still another strategy may need to be employed.

In addition, the use of at least one observational measure is highly recommended. To develop an observation instrument for measuring affect, the evaluator must ask "What behaviors or actions would be likely indicators of the attitude we want to measure?" The starting point for answering this question is to find a term which summarizes the desired attitude, such as "tolerance," "cooperation," "appreciation" and so on. Then, list the observable activities, performances, or behaviors that a person who holds the attitude would be expected to exhibit. Once refined, such a list comprises an observational sheet which can be used as a guide for inferring the extent to which the desired attitudes are present. For example, in an evaluation of a reading program for 2nd & 3rd graders, one of the evaluation questions had to do with "increased desire to read for fun." The measures included examination of the number of children checking out books and the number of books checked out of the library during weekly class visits. Library records were examined before, during and after the program. The children were also administered an anonymous questionnaire concerning their feelings about optional reading and books at the

beginning and end of the program. In addition, the evaluator observed the children 3 times to note how many looked at or read books as opposed to doing other activities during their free periods.

2. *Implementation Questions*

Measurement devices for program implementation should help answer the question, "What does the program look like in operation?" Existing records, such as lesson plans, minutes from meetings, administrative documents, can paint a picture of program operation. It also is quite possible that classroom teachers or other program staff maintain records which can be used as a source of descriptive information on instructional processes. For instance, individualized instructional programs often rely on records to track student progress. Tabulating such data could indicate which program components were used, by whom, and with what results.

Observing and recording the events of key program sessions represents another implementation strategy. A thorough implementation effort could even include photography, tape recordings, and a collection of representative samples of student work.

Data not provided through records or observations can be obtained by using one or more report methods. A word of caution about using only questionnaires and interviews to document a program should be mentioned. Keep in mind that data collected through these two types of instruments are subject to the respondents' selective remembering (i.e., recalling only the "worst" or "best" part of a session), plus biases in favor or against particular aspects of a program operation. Accurate records and controlled observations are therefore generally more useful measurement strategies

for recording implementation factors.

3. *Cost Questions -- A Special Case*

Measuring the "cost" of a program is both a challenging and controversial effort. For example, even sophisticated cost analyses performed by "blue ribbon" teams of economists, accountants, and management scientists often are challenged by groups having a stake in the program under evaluation. This state of affairs stems from the fact that there is no universally accepted definition or set of procedures for investigating cost, and the issue frequently becomes, in the final analysis, a matter of judgement. Ordinarily, it is the legitimate province of decision-makers (or those who commissioned the evaluation) to make judgements about the way in which data the evaluator collected ought to be interpreted, as different analytic schemes such as cost benefit analysis, benefits foregone, etc., result in different conclusions. The evaluator's main responsibility, unless otherwise charged, then is to provide objective and credible data to those people who will, in turn, make educational decisions. Therefore, evaluators should be prepared to carry out limited but credible cost investigations that reflect the stages of the entire "life cycle" of a program. These stages include:

- planning stage -- refining proposed program goals, selecting materials, instructional strategies, etc.
- start-up stage -- what is done to install the program for the first tim
- operational stage -- while the program is running as expected after the "bugs" have been worked out

What follows is one approach to measuring cost that evaluators may wish to apply to small scale evaluations. It is, though, by no

means the only available approach.

Just as the outcomes of instruction must be operationally defined before they can be measured, so must the concept of "cost" be precisely defined. Usually this means breaking down the general term "cost" into more specific categories which, in turn, represent the units for cost analyses. One sample set of categories might include the following:

- personnel — teachers, administrators, aides, etc.
- space — classrooms, offices, etc.
- equipment and supplies — typewriters, buses, computers, paper, etc.

It is quite possible that decision-makers may not be concerned with all three categories; perhaps they are interested only in the use of space in a newly constructed modular classroom or the rate at which supplies must be replenished. Evaluators should establish the "target" categories for cost analysis as early as possible in the course of the evaluation. For example, data could be collected using the following cost indicators:

A. Dollar expenditures including salaries, price paid for equipment, heating, installation of facilities, etc. Excluded from dollar expenditures should be "fixed costs" which exist regardless of the program. Fixed costs might be, for instance, administrative salaries, normal building maintenence, textbooks that are used in both new and old programs, etc.

B. Costs measureable in dollars that are not actual expenditures. This includes "hidden" costs like depreciation of equipment, fringe benefits for staff members hired exclusively for the program, estimates on the cost of space used solely for program instruction, etc.

C. Time spent on the program as well as time normally devoted elsewhere but "given up" because of the program. This included the time spent by teachers, students, administrators, secretaries and others associated with program activities.

D. Costs that cannot be quantified, such as staff morale, community attitude, student satisfaction.

A relatively straightforward strategy for selecting measurement devices to collect cost data can involve displaying the "targets" of cost investigation (categories) against different indicators of cost, as illustrated in the chart below. Only some of the squares have been filled in to serve as examples of measurement devices that might be employed.

Chart 2

Cost Indicators / Categories	A Actual dollar expenditure	B Calculated dollar expenditure	C Time	D Non-quantitativ costs
1. Personnel			A "log" of time spent on planning lessons	Interviews with aides to assess morale
2. Space		Records of utility bills, school property taxes, maintenance costs		
3. Equipment/ supplies	Invoices of payment for materials, checklist of wear and tear and upkeep			Questionnaire t assess teachers and students' satisfaction with materials and equipment

4. *Research Questions*

Since research questions usually include a planned variation in the program processes, these questions most often utilize tests as the measurement device. Therefore, the concerns are similar to those for measuring cognitive and psychomotor outcomes which were discussed in Section H.

5. *Unanticipated Outcomes*

In addition to the expected program outcomes, often the program may have unanticipated outcomes or effects. These unanticipated outcomes may be as important as planned outcomes. They are vital in validating a program so the evaluator needs to uncover and measure them.

One way to systematically measure unplanned effects of a program is to ask individuals, using questionnaires and interviews, if they have reasonable evidence of any other outcomes, either negative or positive, that they would attribute to the program. People who may reveal important information include both the people directly in the program, e.g., teacher, students, as well as peripheral people, e.g., other teachers, parents.

Planning an Efficient Measurement Strategy

It is quite likely that many of the instruments selected to answer the evaluation questions can be adapted to address more than one evaluation concern. In fact, as an overall measurement strategy, it is helpful to select instruments that can be used to answer more than one evaluation question at a time. Fewer instruments will minimize the potential problem of overtesting. An efficiently planned measurement strategy will also make more economical use of available resources (both time and money).

For example, a single questionnaire can serve multiple purposes: it can help provide implementation information for a program by asking teachers to provide information about activities, materials, and schedules; it also can assist with identifying unplanned effects by including open-ended questions such as:

"Do you feel the _____ program has affected students' learning in ways you did not expect? Please give specific examples."

"Do you feel the _____ program has affected the school's normal operation in ways you did not anticipate? Please give specific examples."

The same questionnaire could also investigate concerns about cost by asking respondents about wear-and-tear on materials or equipment, the amount of time spent on the program which ordinarily is devoted to other duties, etc. In most instances, research questions can be addressed through the use of instruments addressing outcome and implementation questions. In all, the evaluator's aim should be to plan a data collection strategy that is comprehensive as well as efficient.

SECTION I

Exercise 6

Developing Measurement Devices

In this exercise you will have the opportunity to try your hand at developing a measurement device to answer an evaluation question from the Experience Based Business Education Program.

Working as a team you will develop a brief questionnaire, a short interview, or an observation schedule to answer the following evaluation question:

How much instruction are students receiving outside the classroom as interns in the local businesses?

The respondent groups will be the business people for the questionnaire and interview, and the students and business people for the observation.

Write at least 5 interview, questionnaire or observation items. An example item is given for each measure.

Please wait for further instructions from the workshop leader before beginning the exercise.

Developing Measurement Devices

Exercise 6

Evaluation Question: How much instruction are students receiving outside the classroom as interns in the local businesses?

Interview

Respondent group: Business people

Items:

1) Did the student have sufficient background knowledge of business practices at the beginning of the internship? If not, in what specific areas is the student deficient?

Questionnaire

Respondent group: Business people

Items:

1) What percentage of time, in the course of a day, were the students involved directly with the customers?

— up to 25%

— 25 to 50%

— 50 to 70%

— 75 - 100 %

Observation

Group(s) observed: Students and Business People

Behaviors:

1) How many times was a student directly involved in a transaction with a customer?

BUILDING IN THE EVALUATION PLAN

SECTION J

EVALUATION DESIGN AND ANALYSIS

SECTION J

EVALUATION DESIGN AND ANALYSIS

J. Design Analysis

1. Design
2. Analysis
3. Exercise 7

Design

In the last section you developed measurement devices like those Peter Slate may develop to answer each of the evaluation questions for the Experience-Based Business Education program. Now, a design must be planned for each measure. A design is a plan for collecting information about a program and the effects of a program. This plan includes target audiences that will be involved as well as the time and manner in which the data from the measure will be collected. In other words, the "who" and "when" of the measurement device. The evaluator must describe the measures and the data collection plan in the design section of the program plan. Designs are used in both formative and summative evaluations, although the specific designs evaluators employ during a formative and summative evaluation may differ.

Formative Evaluation

As mentioned in previous sections of the workshop, a formative evaluation is specifically designed for use when a program is still malleable. That is, the program is still being tried out and refined. The information this evaluation provides tells decision makers what the strengths and weaknesses are in a program and indicates those areas in need of revision. Formative evaluation is of particular use for the evaluation of a new program or a program that is being used for the first time in a district. For a new program, the formative evaluation serves to examine the program component by component. This frequent assessment supplies information regarding the strength of each

component.

The formative evaluator frequently collects information about the students' progress on en-route skills, terminal objectives and data about student attitudes toward the program. Information is also gathered about the implementation of the program. For example, is the program administrator following the plan for implementing the program as outlined by the program planners, and, if not, in what way is the administration of the program different? Are teachers using the program's instructional procedures? Basically, a formative evaluation fully documents a program using frequent assessment and observations to locate components of the program that need revision.

Summative Evaluation

A summative evaluation is appropriate for a program that has been established, i.e., tried out and tested for a period of time in a district. During this time, the program problems, if any, presumably have been solved. The decisions made using the summative evaluation usually include whether to expand, use, or discard the program based on an objective judgment of the overall effectiveness of the program.

The data collected for this evaluation has among its major concerns measuring the level of achievement of the entire program's terminal objectives. These measurements are taken at major intervals during the program and at the end of the program and are often compared with an alternative program. If no other program is available for comparison the summative measurements may be compared to established criteria and the pre-test measures to determine what improvement may be attributed to the effects of the program. Implementation information is also collected during a summative evaluation. If the results of the summative evaluation show insignificant or

no gains, implementation information may be able to demonstrate that the fault is not due to the program itself but to the manner in which it was given. It may show that some component of the program that was instrumental in teaching terminal objectives was weak and interfered with the effectiveness of subsequent components. The interpretation of summative evaluation measurements run the risk of being inaccurate when not accompanied with information gathered during the progress of a program.

Design Plans

Traditionally, design plans are formulated for testing student achievement, the measurement strategy most often used in answering achievement outcome questions and research questions. Designs for testing are usually more complex than those for other measures. Therefore, much of this section will discuss evaluation designs for tests. However, designs involving other measurement devices will also be discussed later in the section.

I. TESTS
II. REPORTS
III. OBSERVATIONS
IV. RECORDS

The designs for achievement testing in formative and summative evaluations differ according to three dimensions that describe the who and when of the measure: (1) comparison groups; (2) randomization; (3) pretests. Generally, program monitoring in formative evaluation does not require the use of control groups or comparison groups as does a summative evaluation. When a program is being modified as it occurs throughout the year, it is difficult to compare it to another "fixed" program. If the need for com-

parison is warranted, however, or if pilot tests within a program are desired, the issues of control groups, randomizing and selecting comparison groups as discussed in this section apply. These three dimensions are outlined for each design in the chart below. A brief discussion of each factor follows the chart. Clearly this chart does not give an exhaustive list of possible evaluation designs for testing achievement. An evaluator may find that an evaluation situation does not fit the criteria for any of the designs listed, and will need to set up another design. Other designs can be found in almost any statistical text book.

Name of Design	Control Group?	Random Assignment?	Pretest?
Formative Designs:			
1. Pretest-Posttest	no	no	yes
2. Posttest only	no	no	no
3. Time Series, Case Study	no	no	yes
Summative Designs:			
1. Pretest-posttest control group	yes	yes	yes
2. Posttest only control group	yes	yes	no
3. Pretest-posttest non-equivalent control group	yes	no	yes
4. Time Series	no	no	yes

Control Group and Randomization

Many evaluation designs, especially in summative evaluations, require assigning students to groups so that comparisons can be made between experimental groups and comparison or control groups. Students in the experimental group receive the full treatment of the program being evaluated. Students in the comparison group participate in an

evaluated. Students in the comparison group participate in an alternative program, often the regular or traditional program. To ensure that the student population of both groups is as equal as possible, a process of random assignment is used. In randomization or random selection, each student has an equal chance of being in any control or experimental group.

The major advantage of random selection is that it provides a cross section of student ability, experience, and maturity. Since randomization assumes that any existing variations in student achievement are evenly distributed throughout the groups, the success of a program cannot be attributed to any consistent bias between the groups. Arguments to which non-randomized procedures may be open include: "The kids in Program X were smarter" or "Program X shows better results because many of the parents of the children in that group were scientists" or even "The control group started at a lower level than the experimental group and so it naturally ended up at a correspondingly lower level." All these hypothetical criticisms illustrate difficulties that could be avoided by implementing a random selection procedure.

Of course, given school constraints, an evaluator cannot always control these factors. For example, often intact classes have to be used for a program thereby eliminating the possibility of random assignment to groups. When a comparison group is made of people who were not randomly assigned, it is called a non-equivalent comparison group. With non-equivalent groups, an attempt is made to have the groups compared be as similar as possible with respect to important factors such as ability.

Pretest

Another consideration for a test design should be: Is a pretest being given? A pretest gives the evaluator a good measure of student

knowledge prior to the program. It ensures that the groups cannot already perform the terminal objective. If they can, the objective of the program obviously must be re-examined. If a comparison group is to be used, a pretest also serves the important function of checking that both groups are fairly equivalent with respect to entering knowledge of the program objectives and thereby comparable.

Another method of evaluation does away with the unit or program pretest and uses only the unit or program posttest. This procedure usually occurs when there is some reason a pretest cannot be given. The program posttest is then compared to posttest measures taken on another program (for example, an SRA Language Kit) whose unit or program objectives are similar. For instance, if a unit is concerned with identifying pronouns, the posttest on the pronoun unit could be compared with the posttest results of the similar SRA unit. Note that when using this method, it is important to make sure that the objectives being measured by both posttests are the same.

In a formative evaluation, one of the best designs is to administer a pretest and posttest at the beginning and end of the program as well as a series of unit pre and posttests for each unit in the program. In contrast, summative evaluations often include a pretest only at the beginning of the program and posttest at major intervals during the program and at the end of the program. The pretest at the beginning of a program should measure critical enroute skills, a modified form of terminal skills, and entry skills (e.g., if the program involved teaching calculus, one would not expect students to perform problems in calculus, but one would expect some proficiency at performing algebraic computations involved in such problems). The unit pretests may be formal or informal in nature. A formal unit pretest measures the entry skills in the unit, and is a

written criterion test. If there are many units, however, you may run the risk of making students "test weary" by giving too many tests. Therefore informal pre-tests, which can be conducted orally, can be used.

The unit posttests provide information regarding the effect of each unit on learning gain. In a formative evaluation, the posttest results should be compared to the unit pretest results. In this way strong and weak units in the program can be targeted.

Time Series Design

Since a time series design differs from conventional designs using pretests and posttests, it is discussed in more detail. The time series design is often most practical to use in the school setting when it may be impossible to randomly assign groups or to find a comparison group. The students in the program are used as their own control group. The same measurement is made on the same group of students at regular intervals at least three times before and three times after the program. By observing if program X seems to affect the trend in the results, it is possible to determine whether program X appears to have had an impact on the outcome measure.

Yet another design, a variation of the time series design, may be useful especially in formative evaluation in letting the evaluator know when a program needs revision. This design, however, instead of taking a series of measurements before and after a treatment to indicate a trend, uses a series of measurements during the treatment to indicate if and when a new trend begins. For example, a fifth grade teacher regularly gave speeded tests on multiplication tables, one objective of the fifth grade math program. It was found that very few of the students were achieving the objective. An intervention

was suggested which involved bringing a group of junior high students to the fifth grade classroom each day. Each junior high student spent 15 minutes drilling one student, then 15 minutes drilling another student on the multiplication table. The fifth grade teacher continued giving the speeded tests during the intervention and noticed that after two weeks there was no remarkable progressive change in the test scores. It was suggested that the intervention time be increased to 20 minutes daily. After one week, an increase in the test scores was noted. This trend continued until the objective was achieved, at which time the intervention was stopped. Continued speeded test measurements showed that the improved test scores remained constant.

I. TESTS

II. REPORTS
III. OBSERVATIONS
IV. RECORDS

The use of reports, observations and records requires a design different from those used for tests, though the evaluator is still concerned with the "who" and "when" of the measure. Unlike design plans for tests, other measurement devices do not necessarily require information about comparison groups, randomization, or pretests. Instead, the evaluator usually needs to decide only what groups of people (e.g., students, teachers, parents) should be given the measure to yield the most information about the evaluation question, and when the measures should be administered. When comparison groups are used, one needs to schedule observations or collect records at the same (or roughly equivalent) times.

Analysis

Although data analysis is performed as the program is running (formative evaluation) or after major portions of the program are completed

(summative evaluation), it should be considered during the evaluation planning. Design and analysis should be considered at the same time since the analysis that can be performed is dictated by the evaluation design. Both should be described in the written program plan.

In this section, three categories of statistical analysis will be discussed: descriptive statistics, tests of differences, and tests of relationships. This workshop will discuss only the more common and simple methods of data analysis. The section is only intended as an overview of the approach and is not meant as a complete course in analyzing data. More sophisticated and complete methods are given in statistics texts or courses.

1) Descriptive Statistics. The evaluator should plan to perform descriptive statistics for all the data that will be collected. Performing descriptive statistics means compiling scores (e.g., number correct on tests, responses from questionnaires) and reducing this raw data into some form that can be easily comprehended by decision makers.

One of the first questions the evaluator may ask is, "What was the average score?" So the evaluator should plan to obtain this information by calculating the mean, median, or mode. Each may give useful information and is easy to calculate. The mean is the statistically average score, the mode is the most frequent score, and the median is the score above and below which 50% of the students scored.

Another useful statistic, the standard deviation, shows how much the scores are spread out around the mean or how much variability there is in the scores. The larger the standard deviation, the more

spread out are the scores.

Scores on pretest A	Scores on posttest B
95	97
90	93
83	86
76	84
71	80
65	79
59	78
54	76
49	72
44	68
686	813
$\overline{X}$ = 68.6	$\overline{X}$ = 81.3
S.D.= 17.4	S.D.= 8.95
% = 67%	% = 81%

With the above information, the evaluator will have a good idea how the scores "look." It will be important to convey this information clearly to those reading the evaluation report. A widely recommended procedure for presenting data is to display the information in the form of tables and graphs. These provide an effective visual image of the data. Therefore, during program planning the evaluator should consider what tables and graphs will be used. These will usually show descriptive information (e.g., scores, averages, percentages).

2) Test of Differences. Where two or more programs are being compared the evaluator will usually want to compare the pre-program and post-program performance of the groups. If the design calls for such comparisons, the evaluator should plan to perform a test of differences. Invariably there will be some difference in scores. It is highly unlikely that groups will have identical scores on any two measures. Therefore, the evaluator can safely assume that the decision-makers will not be particularly interested in the fact that there was a difference. That is, they will want to know

if the difference is a result of the program or if it is a fluke difference that might be the result of chance alone. To answer this question, the evaluator can perform tests of statistical and practical significance. In formative evaluation the decision makers will likely be interested in knowing where in the program differences occurred, so the evaluator will probably run tests of differences on interim measures (i.e., unit tests). In contrast, summative evaluators are usually interested in the overall effectiveness of the program as compared with another program so test of differences would be run not only on within group scores but also between group scores.

3) Tests of Relationships. These measures (the most common of which is the correlation coefficient) show the relationship and strength of relationship between two factors. For example, an evaluator may wish to know if teachers' attitude toward a new program was related to the accuracy of their record keeping. The evaluator may hypothesize that a more positive attitude toward the program results in better record keeping. The evaluator would administer a measure such as a questionnaire or interview to determine the teachers' attitudes. The "scores" from the measure would be correlated with the amount of record keeping.

The following three examples of evaluation designs and analysis plans were developed for the summative evaluation of a counseling program. The plans indicate the who and when of the measures the evaluator selected to answer in an implementation, cost, and outcome question, and the analysis to be performed on the data.

Counseling Outcome Question

Can students who have completed the counseling program identify concepts of time management better than students who did not take the program?

Measures

A. A matching test

Evaluation Design Questions

1) Who will be measured?	Students in the counseling program and a control group of students who have not yet taken the program
2) How will the groups be constitued?	50 students will be randomly selected from those students enrolled in the counseling program. 50 students who have not taken the program will be randomly selected from students on the waiting list.
3) When will they be measured?	Before the beginning of the program and at the completion of the program

Recommended Evaluation Design

Compare the pre- and posttest performance of students assigned to take the course with the performance of the comparison group made up of students who are on the waiting list for next semester's scheduled program. Administer tests to control students in their homeroom on the same days as the tests are given to the experimental group in the counseling session.

Analysis

Measure A - 1) Calculate descriptive statistics (averages, percentages) and prepare charts displaying the statistics.
2) Calculate test of differences and present results in a figure.

Counseling Implementation Question

Are the counselors implementing the instruction of each component of the program as prescribed in the program guides?

Measures

A. Observations
B. Records (Counselor Logs)

Evaluation Design Questions

1)	Who will be observed?	All counselors conducting counseling sessions
2)	When will the observation take place?	At the 2nd, 4th, 5th, 6th, 8th, and 10th sessions
3)	Whose records will be checked?	The counselors
4)	When will the records be checked?	At the conclusion of the last counseling session

Recommended Evaluation Design

Measure A - Six counseling sessions will be monitored by two observers (same observers in each session) equipped with an observation guide used to ascertain if the treatment for that session is being given as prescribed.

Measure B - After the last counseling session, counselor logs will be checked for any indications of deviation from the prescribed treatment.

Analysis

Measure A and Measure B - Calculate descriptive statistics (averages, percentages) and prepare charts displaying the statistics.

Counseling Cost Question

What is the cost of the counseling program in terms of counselor time, for counseling and management functions?

Measures

A. Reports (Interviews)
B. Records (Time Sheets)

Evaluation Design Questions

1) Who will be interviewed?	Counselors
2) When (how often) will the interviews take place?	At the end of the last counseling session of the semester
3) When will the records be examined?	At the end of the last counseling session

Recommended Evaluation Design

Measure A - At the end of the last counseling session, interview counselors to obtain information as to how much time counselors spent interviewing teachers, collecting class records, interviewing clerks, collecting student data from the school office, and arranging with Duplication Services for the production of necessary forms.

Measure B - At the end of the last counseling session, examine counselor time sheets for total time spent by counselors in the group counseling sessions.

Analysis

Measure A and Measure B - Calculate descriptive statistics (averages, percentages) and prepare charts displaying the statistics.

Case Study of An Evaluation

The following is an example of the use of design plans and analysis in real life.

Restrooms and Coffee Cups - a Tale of Two Sexes

Ezekiel Thumpkins, President of J. Walter Thumpkins Advertising, was persuaded by his son-in-law Matthias - whom he had hired in the mailroom, but who aspired to a vice-presidency - that his employees were wasting too much company time in the restroom. Ezekiel had long been irritated by his belief that secretaries spent longer than than male employees ("Those women do nothing but go in there and jabber to each other") but now he began to wonder if the men were perhaps also guilty of wasting his money talking about baseball scores (Ezekiel had spent his boyhood in England where he had learned to play cricket; he despised baseball). Matthias also pointed out to him that there was probably a connection between coffee consumption and time spent in the restroom, and that, at the current price of coffee, coffee time and restroom time were placing the company in financial jeopardy.

Ezekiel's second son-in-law - who also worked in the mailroom - aspired to the same vice-presidency as the first (actually, the two did not care; it was their wives who were pushing them up the ladder) - reminded Ezekiel that he had taken a course in statistics in college and therefore was capable of designing an experiment to see what might be done about the problem, and just incidentially steal a march on Matthias. Joe - his real name was Obadiah, but no one could pronounce it, so

they called him Joe - suggested a time series design. For three weeks he had Malvinia Thompson (she was a maiden aunt of the founder J. Walter Thumpkins and had worked for the firm for 55 years) clock both the male and female employees as they went in and out of the restrooms. This was something she had done for years anyway. It was she who first complained about the length of time "those young snips" spent away from their desks. After one week, Joe instituted his program. He hypothesized that if he blew cigar smoke into the women's restroom and cheap perfume into the men's restroom the employees would spend less time there. Joe instituted these conditions for three weeks, then he collected the data and prepared to analyze it. He licked his pencils and got out his statistics books.

Here is the report which he wrote for the Board of Directors:

The program to increase staff efficiency by reducing "down time" in the restrooms has been a smashing success. As will be seen from the table below, changing the air conditions in the men's and women's restroom reduced down time in restrooms.

Men in Restroom
N = 56

	Baseline			3 Weeks Post		t-Test
	N	$\bar{x}$	sd	$\bar{x}$	sd	
Time in Min	12	7.32	2.34	1.23	.56	t = 3.49*

Women in Restroom
N = 62

	Baseline			3 Weeks Post		t-test
	N	$\bar{x}$	sd	$\bar{x}$	sd	
Time in Min	10	8.05	2.49	2.15	.73	t = 3.70*

* significant

I would estimate that over a five-year period, the savings in employee salaries will be on the order of $245,000.56. It is my recommendation that this program be continued.

I have, in addition, discovered there is in fact a relationship between coffee consumption and time spent in the restroom. I found that the correlation between the two is on the order of .92, which means that more than 80% of the variation in restroom time may be accounted for by coffee consumption (more coffee consumed equals more time in the restroom).

Obadiah Schlumph

At the same meeting, the following memo from the personnel manager was also read to the Board of Directors.

I wish to bring to your attention two serious sources of employee dissatisfaction. It is my belief that these two things are bringing staff morale down seriously, and will seriously affect J. Walter Thumpkins.

1. Restrooms Employees complain that the restrooms smell terribly. Four account executives have left in the past two weeks, taking with them the advertising accounts for Ford Motors, Xerox Corporation, and United States Steel. This loss will cost the company $3,334,456,42.

2. Divorces There has been an alarming increase in divorce actions brought by the spouses of male employees and female employees alike. As near as I can determine, the wives have brought suits because their husbands come home smelling of perfume, and they are claiming infidelity. The husbands of the female employees are claiming the same thing because they say their wives come home smelling of cigar smoke.

J. Brackenbush
Personnel Manager

Section J

Exercise 7

Determining Appropriate Evaluation Design and Analysis

Directions:

In this exercise, only three of the evaluation questions and measures have been pulled out of this year's evaluation plan of the Experience Based Business Education program.

The measures and "who" will receive the measures have been filled in. Working as a team, complete the "when" of the measure. Then use this information to describe an appropriate evaluation design and to choose the appropriate category or categories of data analysis.

Section J

Exercise 7

Determining Appropriate Evaluation Design and Analysis

Outcome Question #2

Are students able to list the responsibilities of both a business person and a consumer in a business transaction?

Measures

A. Criterion-Referenced test

Evaluation Design Questions

Who will be measured? Students in the experienced based business education program.

When will they be measured? _____

_____

_____

_____

Evaluation Design

_____

_____

_____

_____

_____

Analysis

___a. Calculate descriptive statistics and prepare a chart displaying the statistics.

___b. Calculate test of differences and present the results in a figure.

___c. Calculate test of relationship and indicate significant correlations in a table.

___d. The design does not lend itself to analysis. Evaluator should choose another design

Implementation Question #5

Are teachers implementing the program as prescribed in the Teacher's Manual?

Measures

A. Observation

B. Lesson Plans

Evaluation Design Questions

Who will be observed?	Teachers
When will the observations take place?	_____
What lesson plans will be examined?	Teachers
When will the lesson plans be examined?	_____

Evaluation Design

Measure A. _____

_____

_____

_____

Measure B. _____

_____

_____

_____

Analysis

___a. Calculate descriptive statistics and prepare a chart displaying the statistics.

___b. Calculate test of differences and present the results in a figure.

___c. Calculate test of relationship and indicate significant correlations in a table.

___d. The design does not lend itself to analysis. Evaluator should choose another design.

Cost Question #2

What is the effect of the internship on performance in other classes

Measures

A. Questionnaire

B. Test Records

Evaluation Design Questions

Who will receive the questionnaire?	Students
When will the questionnaire be given?	
What test records will be examined?	Student test records from other classes. Records of first 10 weeks (before the internship) and records of last 10 weeks (during internship)
When will the records be examined?	

Evaluation Design

Measure A. _____

Measure B. _____

Analysis

Measure A.

___a. Calculate descriptive statistics and prepare a chart displaying the statistics.

___b. Calculate test of differences and present the results in a figure.

___c. Calculate test of relationship and indicate significant correlations in a table.

___d. The design does not lend itself to analysis. Evaluator should choose another design.

Measure B.

___a. Calculate descriptive statistics and prepare a chart displaying the statistics.

___b. Calculate test of differences and present the results in a figure.

___c. Calculate test of relationship and indicate significant correlations in a table.

___d. The design does not lend itself to analysis. Evaluator should choose another design.

BUILDING IN THE EVALUATION PLAN

SECTION K

WRITTEN PROGRAM PLAN

SECTION K

WRITTEN PROGRAM PLAN

Peter Slate, and the program staff have gone through the process of planning the program and evaluation of the program. Therefore all sections of the written program plan, except the summary, are complete. To review the components of a program plan, an outline which was also included in Section C, is on the next page. This is followed by an exercise designed to give practice in writing a program plan summary.

Outline of a Program Plan

I. Summary

A. Rationale, purposes, justification for the program, brief description of the general goals, procedures, and evaluation.

II. Program Description

A. Instructional Characteristics

a. Program goals

(1) general goals and specific objectives

b. Major program activities

(1) description of activities

(2) sequence of activities

(3) instructional components of each major activity

B. Administrative Characteristics

(1) pupils served by program

(2) resources needed - materials, facilities, equipment, personnel

(3) schedule - when do the program's activities take place, duration of activities

(4) job responsibilities of those implementing the program

III. Evaluation Plan

A. Purpose

(1) general formative or summative information needed by decision makers

(2) specific evaluation questions - outcome, cost, implementation, research questions

B. Design

(1) participants - who will be measured

(2) measures - the tests, questionnaires, etc. to answer each evaluation question

(3) data collection plan - when measures will be gathered for each evaluation question

C. Budget for the evaluation

D. Data Analysis

E. Reporting Evaluation Results

SECTION K

Exercise 8

Model Program Plan

In this exercise refer to the Model Program Plan of the Experienced Based Business Education program for the information necessary to write the summary for the program plan.

Each team member should write a summary one-half to one page in length.

The summary should include brief statements of:

— the program's rationale

— the program's goals

— the general goals, and procedures of the evaluation

SECTION K

Exercise 8

Model Program Plan

Experienced-Based Business Education

Summary

Model Program Plan

I. Summary (to be completed in exercise 7)

Description

II. Program: Experience-Based Consumer Education

A. Instructional Characteristics

a. Program goals

(1) General Goal

To provide students with the practical information needed to operate a business through an experienced-based business education program

(2) Specific Objectives

As a result of their intern experiences in various businesses students will be able to demonstrate how a business person manages resources

Students will demonstrate an understanding of responsible business practices

b. Major program activities

(1) Description of Program Activities

Content is presented through in-class activities and ten week internships in local businesses. Teacher guided discussion and paperbound books with detachable exercises provide additional instruction and feedback

(2) Sequence of Program Activities

Internships will be scheduled in the last ten weeks of the semester. Students will actively participate in local businesses three days per week, for approximately two hours each day. The 10 businessmen participating in the program have each agreed to instruct three student interns each day.

(3) Instructional components of activities

- internships in local businesses - ten businessmen who run small personal businesses will participate in the program. The businesses are two drug stores; two grocery stores; a clothing store; a stationery store; a bookstore; an insurance company; a meat market; and a funeral home. Businessmen will be given instructional guidelines asking them to provide students with
  1) practical experience in interacting with clients
  2) practice in managing and operating a business under supervision (e.g., bookkeeping, ordering merchandise, advertising
- in-class lectures - lectures dealing with operating a business - bookkeeping and client considerations - lectures dealing with advertising, marketing strategies

Administrative Characteristics

1. Pupils. The program is designed for senior high students from tenth to twelfth grade. Fifteen students are enrolled in each class for a total of 30 students.

2. Resources. The program includes a Teacher's Manual, paper bound books for students, unit tests, a program pretest and posttest.

3. Schedule. The program runs for one semester and includes a ten week internship for each student.

4. Job responsibilities. The teachers are responsible for constructing record keeping systems for such things as lesson completion time, unit test scores. The businessmen are asked to follow the instructional guidelines provided them.

III. Evaluation Plan

A. Purpose

The purpose of this evaluation is to investigate the strengths and weaknesses of the Experience-Based Business Education Program so that it may be revised and improved.

The specific evaluation questions are:

Outcome Questions

1. Are students able to manage money and resources to operate a business?

2. Are students able to list the responsibilities of both a business person and a consumer in a business interaction?

Implementation Questions

1. Are the students receiving instruction outside the classroom as interns in the local businesses?
2. Are the local businesses cooperating in the program?
3. Is a one hour orientation meeting being arranged? Is it being attended by parents, businessmen, teachers and students?
4. Are the internships being scheduled so as to not interfere with attendance in other classes?
5. Are teachers implementing the program as prescribed in the Teacher's Manual?

Cost Questions

1. What can be done to keep down the cost of the internships?
2. What is the effect of the internships on other class work?

Research Questions

1. What is the effect of the duration of the internships on achieving course objectives?
2. Does the type of business used for internship affect achievement of course objectives?

Unanticipated Outcome Questions

1. Will the program affect relationship between the business community and the students?
2. Does the course experience affect career choices?

B. Design

(only some of the evaluation questions are provided as examples)

Outcome Question #2: Are students able to list the responsibilities of both a business person and a consumer in a business interaction?

(1) Participants: Students in the program

(2) Measures: criterion-referenced test

(3) Data Collection Plan: Students will be given a pretest on the first day of class and a posttest on the last day of class

Implementation Question #5: Are teachers implementing the program as prescribed in the Teacher's Manual?

(1) Participants: Teachers

(2) Measures: Teacher's logs and observations

(3) Data Collection Plan: Examine teacher's logs at the end of each program unit. In-class observations will be conducted six times during the semester

Cost Question #2: What is the effect of the internships on performance in other classes?

(1) Participants: Students in program. Teachers of students in other classes

(2) Measures: Questionnaires, interviews, test records

(3) Data Collection Plan: Students in the program will receive a questionnaire in the 15th week of the semester. Students' test records in other classes from the first ten weeks (before the internships) will be compared with test per-

formance during the last 10 weeks of the internship.

Research Question #2: Does the type of business used for internships affect achievement of course objectives?

(1) Participants: Students

(2) Measures: Criterion-referenced test used to answer outcome question #2.

(3) Data Collection Plan: See outcome question #2.

Unanticipated Outcome Question #2: Does the course experience affect career choices?

(1) Participants: Students

(2) Measure: Career Performance Questionnaire

(3) Data Collection Plan: Administer questionnaire to students before and after the program.

C. Budget

The evaluation will be supported by funds from the Office of the Assistant Superintendent of Instruction. The amount allocated is $_____ .

D. Data Analysis

(only some of the evaluation questions are provided as examples)

Outcome Question #2: Pretest and Posttest scores will be compared by a test of difference.

Implementation Question #5: Descriptive statistics will be compiled. A chart will show percentage of lessons completed as prescribed in the Teacher's Manual, and percentages of those lessons changed or omitted.

Cost Question #2: Descriptive statistics will be compiled for the questionnaire interview responses; a test of differences will be run between the test record scores of the first 10 weeks and last 10 weeks of the semester.

Research Question #2: See outcome question #2.

Unanticipated Outcome Question #2: A test of relationships will compare the relationship of the ranking order of the career choices between the two questionnaires.

E. Reporting Evaluation Results

Evaluation results of the interim measures will be made to the program planners monthly. Results of end of program measures will be made to program planners one month after end of semester.

SECTION L

FEEDBACK

Section A

Feedback

Exercise 1

Categorizing Evaluation Activities

1. program planning
2. summative evaluation
3. program planning
4. formative evaluation
5. formative evaluation
6. program planning
7. summative evaluation
8. needs assessment
9. program planning
10. needs assessment

Section C

Feedback

Exercise 2

Jobs and Responsibilities in Program Planning

Job Descriptions	Person(s) Responsible
Decide on objectives, instructional materials and teaching strategies, student activities	P
Ensure that objectives are clear, so they can be measured	E
Define program resources (time, money, personnel); constraints and facilitating factors (new equipment); special requirements (legal questions)	P
Determine the extent of evaluator's involvement in program planning	PE
Develop or select tools and procedures for selecting among existing alternative programs	E
Choose a program to meet the District's needs	P
Place restrictions on the running of the program (e.g., random assignment of students)	PE
Specify the purpose of the evaluation	P
Schedule the evaluation	PE
Describe what evidence will be accepted to judge the program's effectiveness	P
Select methods and measures for assessing program success	E
Identify data-collection requirements	E
Plan the data analysis	E
Write the program plan	PE

SECTION D

Feedback

Exercise 3

Program Planning and Evaluation Context

Directions: As you read Documents 1-8 look for the information provided about the scope and context of the evaluation. Fill in only the information indicated on the answer sheet that corresponds to your team identification (A, B or C). Write the number of the document containing the information in the right hand column under the heading Document Source.

I. Program Characteristics	Document Source
What is the subject area?	
business education	*1, 2, 3, 5, 6, 7, 8*
What are the general goals?	
To provide students with information about and practice in managing money and resources to operate a business person and a consumer in a business interaction	*1*
How were the goals established? (Needs assessment, mandated, other)	
needs assessment	*1, 2*
Is there a program selected already?	
no	*2, 8*
Is it in operation?	
no	*2, 8*
Will one be selected or adopted?	
selected	*2, 5*

What kinds of and how many students does it serve? (SES, age, ability, ethnicity, number) — 1,2

inner city high school students,

2 classes each semester

What is the teaching method, style of the program?

experienced-based — 1, 6, 8

II. Program Planning & Evaluation Requirements	Document Source
What is the general purpose of the evaluation?	
implement a program; gather implementation data to improve the program	2
What decisions will be made?	
to modify and improve the program	2
How much time and money are available to plan the program?	
$XX, 3 months	2
How much time is allotted for the evaluation?	
one semester	2
Are there mandated requirements for the evaluation?	
no	2

What kind of evidence is required, desired, and possible? — 2

performance on CRT; testimonials of successful business people

II. School Organizational Structure — Document Source

Who are the decision makers?

Taft, school board — 2

What are their positions in the organization?

superintendent, school board members — 2

Why are they involved in the evaluation? (self-initiated, required)

self-initiated — 2

What are their motives for having the program evaluated

want the program to do well — 3

What is their knowledge of evaluation?

of state assessment — 3

What is their committment to/attitude toward evaluation?

Taft leery; no information on board; Williams positive; Slate open & enthusiastic — 3

Do they agree with the proposed requirements for the evaluation?

yes — 3

How much experience does the planning group have in business education?

Blakely & Crennin - business ed teachers, she for 10 years, he for 1 year, have attended a Business Ed workshop. Martin-Economics Business teacher for 5 years, volunteer on Citizen's Advisory Group of Chamber of Commerce; Taft, Williams, Slate - no experience — 2

How much experience does the group have in evaluation?

not at all — 3

What attitudes do the team members have toward evaluation?

Taft - reserved; Blakely - guarded; Smith - negative; Crennin & Williams - positive with hopes it will turn out negative; Slate - open — 2,3

What is the place of the evaluation in the organization? (local, district, or state level)

district — 1,2

Who does the evaluator report to and work with?

principal, teachers — 2,8

IV. Community Organizational Structure

What involvement with the program do parents businesspeople and other community groups have? *businesspeople will take interns* — 5

What information about the program and evaluation do they have?

none; information is contained in editorials and parent advisory meetings but these are not publicly disseminated — 6,7

What attitudes toward the program do community people have?

some businesspeople for; some against. parents for — 2,4,5,6

Section E: Feedback

Exercise 4: Choosing a Program

FINAL RATINGS OF ALTERNATIVE COURSES

PROGRAM CONSIDERATION	Weight	Program A Rating*	Program A Score	Program B Rating*	Program B Score	Program C Rating*	Program C Score
A. GENERAL CHARACTERISTICS							
1. Goal Agreement	10	2	20	5	50	5	50
2. Cost	7	5	35	3	21	1	7
3. Effectiveness	10	4	40	4	40	2	20
4. Instructional Approach	4	0	-	4	16	5	20
B. INTERNAL QUALITY							
1. INSTRUCTION							
a. Measurable objectives	4	2	8	5	20	1	4
b. Instruction on objectives	10	3	30	5	50	2	20
c. Practice and feedback	6	2	12	5	30	2	12
d. Technical accuracy	4	4	16	5	20	3	12
e. Interesting	4	3	12	4	16	5	20
f. Ease of use	3	5	15	0	-	5	15
2. ASSESSMENT							
a. Frequent checks	6	5	30	2	12	0	-
b. End-of-program	2	5	10	5	10	4	8
3. RECORD KEEPING	1	5	5	0	-	5	5
C. MANAGEMENT							
1. Training and orientation	1	0	-	0	-	5	5
2. Clear procedures	3	0	-	3	9	3	9
3. Other program interference	2	0	-	0	-	0	-
4. Positive student performance on other programs	2	2	4	0	-	0	-
5. Personnel availability	3	3	9	0	-	0	-
6. Space	6	0	-	0	-	0	-
7. Facilities, equipment and materials	6	4	24	3	18	5	30
8. Logistics, communication, and time	6	0	-	0	-	4	24
Subtotal			178		249		237
Total of 6 considerations to be filled in			92		63		24
Final Total			270		312		261
Final Average Rating			2.70		3.12		2.61

*Ratings: 5=Very Desirable; 4=Desirable; 3=Average; 2=Undesirable; 1=Very Undesirable; 0=Not present

Section F

Feedback

Exercise 5

Formulating Evaluation Questions

Outcome Questions

1. Are students able to manage money and resources to operate a business?
2. Are students able to list the responsibilities of both a business person and a consumer in a business interaction?

Implementation Questions

1. Are the students receiving instruction outside the classroom as interns in the local businesses?
2. Are the local businesses cooperating in the program?
3. Is a one hour orientation meeting being arranged? Is it being attended by parents, businessmen, teachers and students?
4. Are the internships being scheduled so as to not interfere with attendance in other classes?

Cost Questions

1. What can be done to keep down the cost of the internships?
2. What is the effect of the internships on other class work?

Research Questions

1. What is the effect of the duration of the internships on achieving course objectives?
2. Does the type of business used for internship affect achievement of course objectives?

Unanticipated Outcome Questions

1. Will the program affect relationship between the business community and the students?
2. Does the course experience effect career choices?

Section I

Feedback

Exercise 6

Developing Measurement Devices

Interview:

Respondent Groups: Business people

Items:

1. Did the student have sufficient background knowledge of business practices at the beginning of the internship? If not, in what specific areas was the student deficit?
2. Do you feel the internship was a valuable way for the student to receive instruction in business education?
3. Did you critique the student while he/she was handling a transaction with a customer?
4. What percentage of time, in the course of a day, was the student usually involved directly with the customers? How did you decide on this amount of time?
5. Did you allow the student to transact business in your absence? If not what reservations did you have?
6. What do you feel was the student's attitude toward your instruction? What indicators did you have of his/her attitude?

Questionnaire:

Respondent Groups: Business people

Items:

1. What percentage of time, in the course of a day, were the students usually involved directly with the customers?

   _____ up to 25%

   _____ 25 - 50%

   _____ 50 - 75%

   _____ 75 - 100%

2. What kinds of transactions, involving customers, did the student conduct directly? (Check all those that apply.)

   _____ customer inquiries

   _____ customer in-person inquiries

   _____ customer telephone inquiries

   _____ purchasing transactions (over the counter, check, cash or credit)

   _____ customer complaints

3. How much of the student's working time was spent receiving instruction from you? (Check the appropriate answer.)

   _____ up to 25%

   _____ 25 - 50%

   _____ 50 - 75%

   _____ 75 - 100%

4. When during the day did this instruction usually occur? Check the appropriate answer(s).

_____ begining of the student's working day

_____ throughout the student's working day

_____ at the end of the student's working day

5. Did you critique the student when he/she was handling a transaction with a customer?

_____ yes _____ no _____ sometimes

6. What was the student's usual attitude toward your instruction? Please rate on a scale from 1 - 5 with 1=highly receptive, and 5=very negative:

1	2	3	4	5
highly receptive				very negative

Observation:

Groups Observed: Business people, students

Behaviors:

1. How many times was a student directly involved in a transaction with a customer?
2. Was the business person present at the time of any student/customer interaction?
3. If the business person was present during the above transaction, did he give feedback during or at the end of the transaction?
4. What type of feedback is given (positive, negative instructional)?
5. How many times does the student seek instruction?
6. What kinds of transactions was the student allowed to conduct?

Section J

Feedback

Exercise 7

Determining Appropriate Evaluation Design and Analysis

Outcome Question #2
Are students able to list the responsibilities of both a business person and a consumer in a business transaction?

Measures
A. Criterion-referenced test

Evaluation Design Questions	
Who will be measured?	Students in the experience based business education program
When will they be measured?	Before the beginning of the program and at the completion of the program.

Evaluation Design
Compare the pre- and posttest performance of students in the experience based business education program. Administer pre-tests to students the first day of the program and posttest on the last day of class.

Analysis

a. Calculate descriptive statistics and prepare a chart displaying the statistics.

b. Calculate test of differences and present the results in a figure.

Implementation Question #5

Are teachers implementing the program as prescribed in the the Teacher's Manual?

Measures

A. Observations

B. Lesson plans

Evaluation Design Questions

Who will be observed?	Teachers
When will the observations take place?	*6 times during the semester*
Whose lesson plans will be examined?	Teachers
When will the lesson plans be examined?	End of each observation

Evaluation Design

Measure A. Each class will be observed *6 times during the semester.*

Measure B. On each day of observation, lesson plans will be reviewed for description of basic instructional sequences and activities.

Analysis

a. Calculate descriptive statistics and prepare a chart displaying the statistics.

Cost Question #2

What is the effect of the internship on performance in other classes?

Measures

A. Questionnaire

B. Test Records

Evaluation Design Questions

Who will receive the questionnaire?	Students
When will the questionnaire be given?	*15th week of semester/ end of semester*
What test records will be examined?	Student test records from other classes. Records of first 10 weeks (before the internship) and records of last 10 weeks (during internship).
When will the records be examined?	At the end of the semester.

Evaluation Design

Measure A. Students will complete a questionnaire in the 15th/last week of the semester.

Measure B. Student Test Records from other classes will be examined. Performance on tests in the first 10 weeks (before the internship) will be compared with test performance from the last 10 weeks (during the internship).

Analysis

Measure A.

a. Calculate descriptive statistics and prepare a chart displaying the statistics.

Measure B.

a. Calculate descriptive statistics and prepare a chart displaying the statistics.

b. Calculate test of differences and present the results in a figure.

Section K

Feedback

Exercise 8

Model Program Plan

I. Summary

The Experience-Based Business Education Program is being initiated at John F. Kennedy. The program's goal is to provide students with information about and practice in managing money and resources to operate a business. A business education program will be selected in response to the results of a need's assessment which revealed that high school students have little knowledge of business practices. An evaluation of the program will provide information to program staff for program improvement. Information for program improvement will be provided by unit pre and post tests, and by an end of a program test that came with the program.